Title
The language of Fiction in The Tale of Genji: Linguistic Representation of Characters' Internal States in Literature

Author
Azuma, Ayano

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The language of Fiction in *The Tale of Genji*:

Linguistic Representation of Characters' Internal States in Literature

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Asian Languages and Cultures

by

Ayano Azuma

2019

ABSTRACT OF THE DISSERTATION

The language of Fiction in *The Tale of Genji*:

Linguistic Representation of Characters' Internal States in Literature

by

Ayano Azuma

Doctor of Philosophy in Asian Languages and Cultures

University of California, Los Angeles, 2019

Professor Shoichi Iwasaki, Chair

This dissertation investigates sentences of narration rendered in the text of *The Tale of Genji* that read as if the reader were experiencing a character's internal states, such as consciousness, perceptions and, emotions on his own. This type of sentence is generally known as 'free indirect speech' and has been used as a technique in European literature since the 19th century (Pascal, 1977). However, these sentences (henceforth sentences of represented internal states: SRIS) are also observed in *The Tale of Genji*, a representative Japanese novel written in the 11th century (known as the Heian period). In the studies of *The Tale of Genji*, extensive research has been conducted on this type of sentence, but the majority of research mostly examined the usage of honorifics by using a qualitative method of close reading, which has left other linguistic elements not fully examined.

Employing a two-pronged method including quantitative analysis and qualitative analysis, the current study attempts to identify linguistic elements that contribute to creating SRIS and to investigate how these linguistic elements are used in this type of sentence as well as the effects it creates in *The Tale of Genji*. The current quantitative analysis indicates that

lexical items associated with subjectivity are statistically proved to be less preferred in sentences of narration, yet they are preferably utilized in SRIS. This finding matches the claim made in previous linguistic studies. That is, subjectivity is the key linguistic quality creating this type of sentence across different languages. In addition, examining specific scenes known as the *kaimami* (peering) in *The Tale of Genji*, the current qualitative analysis reveals that SRIS play an important role and create various effects in context. First, SRIS provide vividness to the text by rendering a direct representation of a character's here-and-now in sentences of narration. Secondly, SRIS make it possible to highlight a piece of information that plays a crucial role in the story as well as to add depth especially to the visual description of a scene.

The current findings and discussion provide a new perspective on the on-going discussion of SRIS in the textual analysis of *The Tale of Genji* by incorporating a quantitative method into literary textual analysis. Furthermore, it also indicates that there is a common linguistic quality, subjectivity, to render this type of sentence across different languages in order to enhance this characteristic phenomenon in literary texts.

The dissertation of Ayano Azuma is approved.

Michael D. Emmerich

Robert S. Kirsner

Hongyin Tao

Shoichi Iwasaki, Committee Chair

University of California, Los Angeles

2019

To my mother

For her endless love, support, and encouragement

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF GLOSS ABBREVIATIONS

ADJ	adjective
ADJ V	adjective verb
ADV	adverb
AUX	auxiliary verb
INTJ	interjection
N	noun
V	verb
V (humble)	humble form of honorific verb
V (polite)	polite form of honorific verb
V (respect)	respect form of honorific verb
P (adverbial)	adverbial particle
P (bound)	bound particle
P (case)	case particle
P (conjunctive)	conjunctive particle
P (final)	final particle
PN	pronoun
PREF	prefix
SUFF	suffix

ACKNOWLEDGMENTS

This dissertation is the very first step to my dream. For completion of this dissertation, I owe a debt of gratitude to many people inside and outside of academia. This dissertation would not have been possible, had it not been for them. They all have helped me grow as a student, a researcher, and a human being.

First and foremost, I would like to express my gratitude to my advisor, Professor Shoichi Iwasaki, who was willing to accept my research. I had become interested in research regarding the intersection between linguistics and literature (literary linguistics) when working on my undergraduate thesis. Looking for a graduate school who might be interested in my research inside and outside of Japan to pursue my research interest, I only received so many disappointing replies. Professor Iwasaki was the first and only professor who showed his interest in my research and encouraged me to apply to a graduate program at UCLA. Without his guidance, support, and constant encouragement, this dissertation would not have been completed.

I would also like to express my gratitude to my dissertation committee members: Professor Hongyin Tao, Professor Michael Emmerich, and Professor Robert Kirsner. I am grateful for Professor Hongyin Tao. He has always been willing to offer personal discussions, provide keen perspectives, especially, on quantitative analysis, and help me sort out issues. I am also indebted to Professor Michael Emmerich for his vast knowledge on *The Tale of Genji* and valuable comments from a literature perspective. Without fail, he has provided me with inspiration and guidance on my findings not only for my dissertation but also for other on-going projects. I also thank Professor Robert Kirsner for his encouragement, kindness, and clarity in his comments and feedback. He has always been willing to offer help to restructure my argument as well as concerned about my well-being.

My gratitude also goes to Professor Olga Yokoyama. Since the first day I took a

seminar with her during my first year of the master's program, she has always been willing to hold personal discussions and help me sharpen my perspective on data analysis. Her guidance, encouragement, and confidence in me have made it possible to come this far. I also thank Professor Michael Marra, who passed away in 2011. Upon my first arrival at UCLA for the master's program in 2008, I received an email from him. He welcomed me with open arms and invited me to have dinner, where we talked about my research interests. His kindness, encouragement, and passion for Japanese literature will always be remembered.

I have received technical support for my quantitative analysis. I am indebted to the Smart Systems Institute, the National University of Singapore for providing me an opportunity to work with their engineering team for my research project.

I extend my gratitude to my friends as well as fellow graduate students: Grigor Ketenchian, Daphne Yu-Hui Lee, Kyoungmi Ha, Yang Ying, Madeline Kubo, Jihyeon Cha, Kirk Kanesaka, Don Lee, Derek Chin, Kayi Ho, and Gabriel McNeill. They have supported me inside and outside of school and helped me keep going throughout my graduate studies. Without their friendship, constant support, and encouragement, this dissertation would not have been possible.

Finally, I would like to thank my family in Japan, who have wholeheartedly supported me to pursue my research interest in the United States. My husband has provided full support, kept me sane on my worst days, and offered the warmest words of encouragement during this long journey. Lastly, I thank my mother, who passed away during my third year of the doctoral program. She taught me commitment, perseverance, and dedication and always let me be who I am. My childhood experiences that she provided me have been the prime inspiration for my research, which will always be deep in my heart. This dissertation is dedicated to her memory.

VITA

2011	M.A., East Asian Studies University of California, Los Angeles
2006	B.A., Foreign Languages (English) Department of English Dokkyo University
2018-2019	Graduate Intern The Smart Systems Institute The National University of Singapore
2013-2015	Graduate Student Researcher Department of Asian Languages and Cultures University of California, Los Angeles
2018 (Winter)	Teaching Fellow Department of Asian Languages and Cultures University of California, Los Angeles
2014-2016	Teaching Associate Department of Asian Languages and Cultures University of California, Los Angeles
2012-2014	Teaching Assistant Department of Asian Languages and Cultures University of California, Los Angeles
2010 (Fall)	Teaching Assistant Department of Asian Languages and Cultures University of California, Los Angeles
2009 (summer)	Teaching Assistant Department of Asian Languages and Cultures University of California, Los Angeles

SELECTED PUBLICATIONS AND PRESENTATIONS

Takeuchi, A. (July, 2019). Qualitative Meets Quantitative: A Textual Analysis Case Study on the Text of *Kaimami* Scenes in *The Tale of Genji*. The Poetics And Linguistics Association (PALA), Liverpool, England.

Takeuchi, A. (July, 2018). Gender is rendered in "crying": A quantitative analysis on descriptions of crying in the poems of *The Tale of Genji*. The Poetics And Linguistics Association (PALA), Birmingham, England.

Takeuchi, A. (July, 2017). Experiencing a scene in the novel "*The Tale of Genji*": A case study of classical Japanese. The Poetics And Linguistics Association (PALA), Philadelphia, USA.

Takeuchi, A. (June, 2015). A linguistic analysis of the poems in *The Tale of Genji*: "crying" shapes gender. Japan Society of Stylistics, Tokyo, Japan.

CHAPTER 1

INTRODUCTION

1.1. Objectives

This dissertation aims to provide a new paradigm of the textual analysis of *The Tale of Genji*. It proposes a new framework for discussing sentences of narration that read as if the reader were experiencing the scene or the feelings of a character on his own (henceforth sentences of represented internal states: SRIS) as well as the existence of the narrator in the tale by analyzing linguistic features that index subjectivity. In traditional textual analyses of the tale, SRIS have been mainly discussed through examining the use of honorifics. However, the current study will investigate several other grammatical features, such as certain auxiliaries and exclamatory expressions, that are associated with subjectivity. Subjectivity has been playing a crucial role in the development of theoretical discussion of the Japanese language both by traditional Japanese linguists, such as Yamada (1908) and Tokieda (1941, 1950), as well as linguists trained in the West, such as Kuno (1973, 1987), Kuroda (1973/ 2014), Akatsuka (1978, 1979), and Iwasaki (1993). This concept will be set as a guiding post for current analysis, relying especially on the seminal work by Banfield (1982) and Kuroda (1973/ 2014), which investigated the role of subjectivity in SRIS across different languages. In addition, this dissertation provides a methodological innovation using two different methods: a qualitative method of close reading which was employed in previous textual analyses, and a largely unexplored quantitative method of corpus linguistics. The corpus linguistic method will allow us to perform a data-driven analysis, which is now made possible thanks to the advancement of computational technology. Based on linguistic data obtained through quantitative analysis, this new approach allows us to further investigate how the common linguistic quality of subjectivity is utilized to create SRIS, which appear in literary works of various languages.

The Tale of Genji, a Japanese literary work written in the 11th century (the Heian period) that is sometimes described as the oldest extant novel, has provided scholars in different disciplines with various challenges for many years, such as hermeneutics, poems, and picture scrolls. Using both quantitative and qualitative approaches, this dissertation examines one aspect of the tale from a linguistic point of view. Specifically, it focuses upon "sentences of represented internal states" (SRIS), which allow readers to experience characters' subjective experiences. In third-person narratives, this type of sentence is rendered as narration, which is supposed to be written from the third-person perspective. Nevertheless, it can also be read from the first-person perspective (Mitani, 1994). This dissertation will first statistically identify the word distribution of lexical items across three different text types (e.g., narration, conversation, and short poem), which allows us to identify the distributional pattern of lexical items associated with subjectivity across these text types. Identifying the word distribution in the tale also enables us to determine lexical items as well as linguistic functions that appear prominently in each text type. Based on the findings from this statistical analysis, the current study determines which lexical items associated with subjectivity to examine and find where in the tale these lexical items are utilized, especially in sentences of narration. The study then turns to qualitative analysis to examine effects that "sentences of represented internal states" (SRIS) create in context.

Due to its characteristic phenomenon, SRIS have attracted attention from many researchers and has been investigated in the field extensively, especially in the 1970s and 1980s. Among the linguistic elements that have intensively been investigated are honorifics (e.g., Mitani, 1994; Takahashi, 1991; Jinno, 2004). Honorifics are utilized to indicate the social status of the participants in a sentence, such as the agent, the patient, and the participant. Since the Heian period was an aristocratic society, honorifics were essential linguistic elements in the language system. However, in sentences of represented internal states, honorific expressions tend to be omitted when necessary (Mitani, 1994). Due to its

unique linguistic behavior, the usage of honorifics has been intensively investigated, which led to crucial discussions in the field, for example, of how to classify sentences of represented internal states (e.g., Nishio, 1952/ 1991; Mitani, 1994; Kai 1980) and how to interpret the existence of the narrator in the tale (e.g., Takahashi, 1991, Mitani, 1991, Jinno, 2004). Non-honorific grammatical elements (e.g., auxiliaries), however, have not yet been researched adequately. As such, this dissertation aims to provide linguistic analysis for the on-going discussion of SRIS, especially non-honorific linguistic elements, found in *The Tale of Genji*.

This dissertation adds a hitherto unexplored quantitative methodology to investigate SRIS rendered in *The Tale of Genji*. Recent developments in computational linguistics together with the availability of large corpus make this process possible. By combining quantitative analysis and qualitative analysis, this study may provide a data-driven perspective to textual analysis of *The Tale of Genji* with empirical linguistic evidence.

The ultimate goal of this dissertation is 1) to find linguistic elements that are preferably used in SRIS, 2) to identify a linguistic characteristic that contributes to creating a unique phenomenon in the text, and 3) to analyze the way these linguistic elements are utilized in the text as well as the effects SRIS create in context. By employing a two-pronged method using quantitative and qualitative analysis, the current research attempts to shed new light on the textual analysis of *The Tale of Genji* from a linguistic point of view. Furthermore, this dissertation aims to contribute to the field of literary linguistics, which is an interdisciplinary field that is dedicated to the research at the interface of linguistics and literature. By incorporating a quantitative analysis into textual analysis through extensive use of the computer, this study hopes to demonstrate how a linguistic approach may benefit textual analysis of literature.

1.2. The Scope of the Study

1.2.1. SRIS (Sentences of Represented Internal States)

SRIS, which are investigated in this dissertation, are also known by various other names, including 'free indirect speech,' 'narrated monologue,' and 'style indirect libre,' to name a few (Prince, 1987/ 2003). Prince describes this type of sentence as follows:

> A type of discourse representing a character's utterances or thoughts. … which is not linguistically derivable from direct discourse or ("normal") tagged indirect discourse … In other words, free indirect discourse is not definable in strictly grammatical terms. (p. 34)

As explained in Prince (1987/ 2003), functionally speaking the 'free indirect speech, 'or SRIS, represents a character's consciousness (feelings, emotions, and thoughts), yet at the same time, it cannot precisely be defined "in strictly grammatical terms." That is, SRIS are unique both functionally and grammatically, and is widely employed in literature written in other languages, such as modern Japanese, English, and German (e.g., Pascal, 1977; Kuroda, 1973/ 2014, Banfield, 1982). In European languages, Pascal (1977) states, it randomly appeared in literary texts since the Middle Ages. However, it was not until the 19th century that writers such as Jane Austen, an English novelist, and Johann Wolfgang von Goethe, a German novelist, began to use this type of sentence as "a prominent and continuous feature in a novel" (p. 34). Gustave Flaubert, a French novelist of the 19th century, was the first writer who realized its stylistic uniqueness and employed it as a technique. As time goes on, Pascal claims, use of this type of sentence has increased in literary texts of European languages, which led to "the evolution of the novel towards the depiction of states of mind, temperament, moods, rather than external actions (p. 34)".

It is believed that Adolf Tobler, a Swiss-German linguist, first "discovered" this unique style in literature (Pascal, 1977). In his article published in 1897, Tobler described this

unique style as "a 'peculiar mixture' of direct and indirect speech (as cited in Pascal, 1977, p. 8).” However, already in 1912, Charles Bally, a French linguist, had already examined this style of sentences. In his article, Bally named this unique sentence as *le style indirect libre*. As Tobler describes, Bally also recognized its unique mixture of direct and indirect speech (Pascal, 1977). While it may have a syntactic structure that is identical with indirect speech, *le style indirect libre* also retains linguistic elements that indicate the subjectivity of a character, such as question, explanation, and personal vocabulary. Bally thus concluded that *le style indirect libre* “enjoys an almost absolute syntactical liberty (as cited in Pascal, 1977, p. 10)” and cannot be defined solely based on grammar. Pascal states that Bally’s analysis stimulated the interest of many scholars, such as literary critics, philosophers, and linguists, and led to an active discussion as well as the growth of stylistic studies in the European languages. As discussed so far, SRIS have been observed across different languages and widely discussed among researchers from various fields. Investigating SRIS rendered in a classical Japanese literary text, this dissertation thus aims not only to introduce an empirical perspective into the textual analysis of *The Tale of Genji* but also to offer further insight into SRIS in general.

1.2.2. Subjectivity

Subjectivity plays a vital role in SRIS. Although it is difficult to define grammatically this unique sentence in literature, subjectivity, as mentioned in the previous section, is a crucial characteristic of SRIS. The notion of subjectivity has recently become a significant topic in linguistics, though the origin of the study can be traced back to philosophic investigations by Descartes, Kant, and Fichte (Narrog, 2012). According to Narrog, in French and Japanese linguistics, subjectivity began to be examined in the 1930s while in English linguistics, it came into the mainstream in the 1970s. John Lyons, whose scholarship is often viewed as the reference point for the study of subjectivity in English linguistics, views

subjectivity as "self-expression in the use of language" (1995, p. 337). Subjectivity, for him, is not something just realized as a certain set of linguistic expressions or encoded in linguistic expressions per se, but something added by the speaker "with various kinds of prosodic and paralinguistic modulation" (Lyons, 1995). Since Lyons, different scholars have investigated subjectivity by employing various approaches.

Recently, subjectivity has been investigated by using empirical data. Adopting the view on subjectivity proposed in Lyons, Traugott and Dasher (2002) say that "expression is neither subjective nor objective in itself; rather the whole utterance and its context determine the degree of subjectivity" (2002, p. 98). In contrast, Iwasaki (2013) claims that subjectivity is "the effect that the speaker self brings to the system of grammar" (p. 287). That is, subjectivity may be realized in language. Analyzing modern Japanese conversations, Iwasaki (1993; 2013) associates subjectivity with various kinds of linguistic expressions and features and classifies them into different categories, such as internal state expressions, mental processes, and epistemic modality. In his analysis, Iwasaki covers a wide range of linguistic expressions, such as particles, exclamation markers, and auxiliaries, claiming that certain linguistic expressions are associated with the speaker's subjectivity and cannot be adequately understood without the notion of the speaker.

In order to facilitate the investigation of SRIS, the current study examines lexical items associated with subjectivity. The importance of subjectivity as well as its unique use in SRIS are also highlighted in Banfield (1982) and Kuroda (1973/ 2014), which will be discussed in Chapter 2. Among various linguistic elements associated with subjectivity, this study examines three linguistic elements associated with subjectivity: auxiliary verbs for speculation, particles for exclamation, and interjection. These lexical items of subjectivity are statistically found to appear prominently in the text, which will be discussed in Chapter 3, and they will be further examined qualitatively in Chapter 4.

Also, out of various scenes in *The Tale of Genji*, the current qualitative analysis

focuses on a specific type of scene known as *kaimami*. *Kaimami*, which conventionally refers to the action of peering where a male character covertly peers at a female character, is very common in narrative literature written in the Heian period, including *The Tale of Genji* (Yoshikai, 2008). *Kaimami* scenes are good places to investigate SRIS for the following reasons: firstly, this unique sentence, Mitani (1994) claims, has a correlation with expressions to indicate "seeing." *Kaimami* scenes tend to begin with a word indicating "seeing," such as *nozoku* (to peer), *miru* (to look at) and *miyu* (to see). That is, there is a high chance of occurrence of SRIS in *kaimami* scenes. Secondly, *kaimami* usually describes not only "what is seen" but also "who sees" (Yoshikai, 2008). In the depiction of the observer, Yohikai states, internal states of a character, such as judgments, feelings, and consciousness, are also rendered.[1] It is thus reasonable to say that there is a higher chance of occurrence of SRIS in *kaimami* scenes. Therefore, the qualitative analysis in this dissertation examines the texts of *kaimami* scenes in *The Tale of Genji*.

1.3.Methodology

This study employs a two-pronged method of quantitative and qualitative analysis. I first conduct a quantitative analysis through extensive use of the computer to analyze a large amount of data, which would otherwise be unfeasible, and then move on to a qualitative analysis on the text of *The Tale of Genji* based on the findings of the quantitative analysis. For the quantitative analysis, I employ a method used in corpus linguistics in order to discover word distribution across the three different text types (narration, conversation, and short poem) in the text, which allows us to examine linguistic functions of lexical items in each text type. Identifying the word distributions in each text type also leads to identifying the distribution of subjective lexical items across the three text types. Based on the findings

[1] Pascal (1977) also claims that in European literature, depiction of characters' internal states correlates with the occurrence of SRIS.

of the quantitative analysis, I conduct a qualitative analysis on the text of *The Tale of Genji*. Focusing on three different linguistic elements that are associated with subjectivity (auxiliary verbs for speculation, particles for exclamation and interjections), this qualitative analysis investigates the way these lexical items of subjectivity are used in the text as well as the effects SRIS create in context.

A corpus linguistic approach is suited for examining the language and texts of classical Japanese. Since classical Japanese is an archaic language, it is impossible to collect new data more than already existing data or to conduct experiments to confirm findings. As such, employing computational technology in analyzing a large amount of data, the current quantitative analysis enables researchers to take an inductive approach to the data and to provide a linguistically systematic investigation and empirical support to literary text analysis. By doing so, this study helps us gain a more in-depth insight into textual analysis of *The Tale of Genji*.

1.4.Organization of the Dissertation

This study first provides an overview of the textual analysis of *The Tale of Genji* and introduces linguistic research concerning SRIS in Chapter 2. Reviewing crucial research on the textual analysis of the tale enables us to perceive the current state of the field and to identify remaining issues. Since the textual analysis of the tale has a long history from the 12th century, it may be beneficial to review crucial studies and highlight achievements of classic commentaries as well as present-day research. Reviewing linguistic approaches to SRIS in modern Japanese literature and English literature, this chapter presents the framework for the current analysis and discusses targeted lexical items for the quantitative analysis (Chapter 3) and the qualitative analysis (Chapter 4).

Chapter 3 presents the quantitative analysis, which is comprised of two parts: the log likelihood ratio (LLR) and concordance plots. This chapter begins with an examination of the

distribution of lexical items across the three different text types (narration, conversation, and short poem) by using a diachronic corpus provided by the National Institute for Japanese language and Linguistics. The log likelihood ratio, the statistical method used in this chapter, makes it possible to identify lexical items that are statistically prominent in each text type. Identifying the word distribution of each text type leads to identifying the distribution of lexical items associated with subjectivity. Based on the findings of LLR analysis, this research then determines targeted linguistic elements associated with subjectivity and create concordance plots to detect where these targeted lexical items appear in sentences of narration in *The Tale of Genji*.

In Chapter 4, the study turns to a qualitative analysis. Focusing on the way these targeted lexical items are utilized and the effect they create in the text, this chapter investigates SRIS appearing in *kaimami* 'peering' scenes in *The Tale of Genji*. *Kaimami* scenes are most important in investigating this type of sentence since the consciousness of a character (consciousness, perceptions, and emotions), as Mitani (1994) claims, is often rendered in *kaimaimi*. This qualitative analysis examines six *kaimami* scenes from five different chapters in the tale: the 3rd chapter (*Utsusemi*), the 5th chapter (*Wakamurasaki*), the 28th chapter (*Nowaki*), the 34th chapter (*Wakana: jo*), and the 46th chapter (*Shiigamoto*). These scenes are the most famous *kaimami* scenes in the tale and play important roles in the story.

Chapter 6 summarizes the major findings in the current research and discusses the implications in terms of how a common linguistic quality (subjectivity in this study) is utilized to create SRIS across languages. It finally makes suggestions for future studies in order to further the textual analysis of *The Tale of Genji* as well as classical Japanese literary works.

CHAPTER 2

LITERATURE REVIEW

In this chapter, I review major literature concerning textual analyses of *The Tale of Genji* and linguistic examinations of sentences of represented internal states (SRIS) in literature. While the textual analysis of *The Tale of Genji* has a long tradition, starting in the 12th century, the study of SRIS in the tale has a much shorter history, starting in the late Edo period (the 19th century). The latter tradition regained its momentum in the 1970s and 1980s. To give a proper treatment of SRIS in *The Tale of Genji*, however, I also need to review the literature of this type of sentence in a broader perspective, because as mentioned in the introduction, this type of sentence appears not only in classical Japanese literature but also in modern literature as well as that written in other languages. Due to its unique characteristics in the text and grammatical behavior, it has attracted the attention of researchers and been investigated by literary scholars as well as linguists from various perspectives. A common thread in linguistic studies reviewed in this chapter is that they examined not particular lexical items but the function of linguistic elements utilized in sentences of represented internal states. Investigating literary texts of different languages, these linguistic studies identified two crucial characteristics of SRIS. First, lexical items associated with subjectivity contribute to creating sentences of represented internal states; secondly, grammatical restriction identified in normal use of language may not always hold in literary texts. In reviewing these linguistic studies, this chapter presents a new perspective from a linguistic point of view into the discussion of sentences of represented internal states in the textual analysis of *The Tale of Genji*.

By reviewing crucial research conducted for the tale, this chapter summarizes the current state of the field, and at the same time, it identifies the remaining issues, especially concerning the nature of SRIS. This chapter contains two major sections. In the first section, I

introduce research on sentences of represented internal states in the western tradition, including both narratology and linguistic studies. By reviewing these studies, this section presents two competing frameworks for literary text analysis and decides which framework is best for this dissertation. The second section provides a brief overview of the textual analysis of *The Tale of Genji*, including both old commentaries and the current state of the textual analysis of the tale. I present the development of the textual analysis of the tale as well as crucial achievements by reviewing major studies concerning sentences of represented internal states, which leads to the identification of remaining issues in the field.

2.1. The Western Tradition

There are two competing frameworks for textual analysis of literary works: the communicational model and the non-communicational model (Banfield, 1982).[2] These two frameworks differ in their view of the essence of language. The framework of the communicational model applies the speaker-hearer relationship to textual analysis, whereas the framework of the non-communicational model does not. These frameworks dictate not only the view on the quality of language in literature but also interpretation of sentences of represented internal states. This section discusses the different perspectives on textual analysis proposed in each framework.

2.1.1. The Framework of the Communicational Model: Narratology

The framework of the communicational model is built on the grounds that the essence of language exists in communication. That is, this framework entails a speaker (narrator) and a hearer (reader) both in the paradigmatic discourse of ordinary communication and in the literary linguistic setting (hereafter, discourse refers to the paradigmatic linguistic

[2] Banfield uses the terms “a communication theory of language” and “a non-communication theory of language."

performative setting of ordinary communication as opposed to the literary linguistic setting). Language in literature, Genette (1972/ 1980) states, is a subcategory of language in discourse, which is a special mode of language with restrictions and limitations not found in regular language use. In order to encompass both language in discourse and language literature, the framework of the communicational model stipulates that to tell is to speak, and all the sentences in literary texts are assigned to a single narrating voice. That is, functioning as the speaker of the text, the narrator is considered to exist throughout the text and is responsible for all the sentences therein. This can be seen in Pascal (1977): "the narrator must provide a language for matters that, for the character, resist verbal formulation" (p.122), and render "a coherent and logical sequence of a character's thoughts" (p. 16). In other words, the narrator is the organizer of the text as well as the locus of the meaning of sentences. As such, an important notion is proposed in the framework of the communicational model: an omniscient narrator. The omniscient narrator is defined in Prince (1987/ 2003) as follows:

> A narrator who knows (practically) everything about the situation and events recounted . . . Such a narrator has an omniscient point of view and tells more than any and all the characters know. (p. 68)

Being aware of everything including characters, places, and events, the omniscient narrator may also take perspectives of any characters in a story. In other words, the omniscient narrator has access to characters' internal states. The existence of the omniscient narrator in texts also governs interpretation of sentences of represented internal states. In the framework of the communicational model, sentences of represented internal states (SRIS), or also known as "free indirect discourse," are defined in Prince (1987/ 2003):

> A type of discourse representing a character's utterances or thoughts. . . which is not linguistically derivable from direct discourse or ("normal") tagged indirect discourse—is usually taken to contain mixed within it markers of two discourse events (a narrator's and a character's), two styles, two languages, two voices, two semantic and axiological system. (p. 34).

According to Prince (1977), this type of sentence is narrated by the narrator and spoken by a character at the same time. That is, two different voices (the narrator's voice and a character's voice) merge, or layer, into one, which is known as the dual voice phenomenon proposed in Pascal (1977). In conclusion, the framework of the communicational model offers to textual analysis on literature a whole system, which also encompasses both languages in discourse and literature.

2.1.2. The Framework of the Non-communicational Model: Linguistic Perspective

Literary texts are investigated not only by literature scholars but also by linguists. This section thus presents linguistic research on the textual analysis of literature concerning sentences of represented internal states: Banfield (1982) and Kuroda (1973, 1974, 1975). Banfield investigated mainly English in reference to French and German while Kuroda investigated modern Japanese. Examining linguistic features used in sentences of represented internal states rendered in different languages, they both found that linguistic elements associated with subjectivity contribute to creating this type of sentence. While investigating SRIS (sentences of represented internal states), they demonstrated that lexical items associated with subjectivity show unique linguistic behavior. That is, grammatical restrictions observed in an ordinary communicative setting may not hold, which leads Banfield and Kuroda to conclude the language of literature, Banfield and Kuroda claim, should be considered to be distinct from the language of discourse of ordinary communication and examined on its own right.

2.1.2.1. Banfield

Banfield examined third person narratives mainly of English literature and syntactically defined the language of literature in comparison with the language of discourse. The language of literature, she claims, behaves syntactically differently from that of discourse; therefore, it should be considered to be a distinct category in the system of language. In particular, the language of literature has unique sentences that completely exclude any expressions or constructions associated with the communicative act. In literary texts, Banfield suggests, it is important to differentiate between the speaker and the subject of consciousness (henceforth SELF following Banfield). The speaker is someone who speaks while the SELF is someone to whom a subjectivity/ point of view is attributed. While in discourse the speaker and the SELF are identical, in literature they may be differentiated especially in third person narratives. The separation between the speaker and the SELF creates two types of sentences unique to literature: sentences of narration and sentences of represented speech, thought, and consciousness. Banfield asserts that these unique sentences are considered to be 'narratorless' sentences (mainly in third person narratives) and constitute the narratives of literature.

Focusing on these two types of sentences unique to literature, Banfield presented syntactical differences between the language of discourse and the language of literature. The first type of narratorless sentence is the sentence of represented speech, thought, and consciousness, which may be read as if the sentence were rendered in the first-person. Examining sentences in English and French literature, Banfield finds that these sentences contain expressive forms, which are associated with subjectivity (e.g., exclamations, verbless exclamatory constructions, and repetitions) yet do not contain any forms that are associated with the communicative acts (e.g., direct address, addressee-oriented phrases and adverbials, and the second person). That is, in these represented sentences, there exists a SELF, who

expresses his own subjectivity, yet a speaker, who performs a communicative illocutionary act, does not exist. This type of sentence, Banfield states, is not spoken; the subjectivity of the SELF is directly revealed without any intervention of the speaker. Moreover, she claims, this feature is not compatible with either direct or indirect quotations, both of which are employed for reporting other people's words in discourse. In direct quotations, the proposition, subjectivity, and communicative features are all rendered, whereas in indirect quotations, only the proposition is rendered. In addition, there is a lack of present tense in the represented sentence. Since present tense represents the moment of utterance in discourse, Banfield argues that the lack of present tense also indicates the absence of the moment of utterance in represented sentences. However, deictic expressions, such as now, tomorrow, and today, may be rendered in represented speech. Banfield provides an example from *The Years* written by Virginia Woolf.

> Leaning out of the window side by side the two women watched the man … Now he threw away his cigarette.
>
> (reprinted from Banfield, 1982, p. 96)

In the example sentence above, she claims that the deixis NOW is used with past tense since it is associated with SELF at a moment of consciousness. However, it lacks the present tense since it is associated with a moment of communication. Providing grammatical evidence, Banfield claims that not the speaker but the SELF is realized in the represented sentence. That is, this type of sentence is narratorless, which is not spoken yet expressed.

Another type of narratorless sentence is pure narration. Pure narration, Banfield states, also occurs only in literary contexts and constitutes an essential part of literary narratives. For example, in French, there is a distinctive marker of pure narration: the past tense form *aorist*. In French, there are two past tense forms: *passé composé* and *aorist*. While

the *passé composé* is used in discourse, the *aorist* is used in literature. Examining sentences in which *aorist* is used, Banfield recognizes that *aorist* does not co-occur with deixis (NOW, THEN, HERE, THERE) or any features of discourse, such as direct address and addressee-oriented adverbs. That is, pure narration excludes any forms associated with the communicative act. Therefore, Banfield claims, it is clear that sentences of pure narration are restricted to literary context. In English, however, there is no equivalent linguistic form to *aorist*. Banfield states that simple past in English is the closest equivalent to *aorist* in French. When simple past is employed for pure narration, it does not co-occur with either deixis or any forms associated with the communicative act, just like *aorist* in French. Example sentences are provided below.

(1) She saw the moon.

(2) She now saw the moon.

(reprinted from Banfield, 1982, p. 157)

Both of the sentences are in the simple past. However, Banfield claims that the sentence (1) may be translated in *aorist*; however, the sentence (2) cannot be translated in *aorist* due to the existence of the deixis NOW. This language behavior suggests that there exists pure narration in literary texts of English as well, which is syntactically different from the language of discourse. In sum, these two types of sentences unique to literature (pure narration and represented speech, thought, and consciousness) exclude any forms or constructions associated with the communicative act. Investigating syntactic features in narratives of literature, Banfield concludes that the language of literature should be recognized as a distinct category from the language of discourse and examined in its own right.

2.1.2.2. Kuroda

Focusing on modern Japanese, Kuroda also claimed the grammatical uniqueness of sentences of represented internal states. Narratives of literature are usually categorized into the first-person narrative or the third-person narrative; however, based on differences in grammatical behavior, Kuroda (1973/ 2014) classified narratives of literature into two groups: reportive style and nonreportive style. In the reportive style, which is usually employed in first-person narratives, conventional grammatical restriction holds, whereas in the nonreportive style, which is only employed in third-person narratives, grammatical restriction may not hold. In other words, the reportive style is used in the paradigmatic linguistic performative setting, or the discourse setting, in which the narrator "I" narrates a story from his own point of view. In an extended sense, Kuroda explains, included in the reportive style are non-first-person narratives where the effaced narrator narrates a story strictly from his point of view. That is, the narrator is not omniscient but omnipresent and narrates none of the characters' inner states. On the other hand, the nonreportive style, which is unique to literature, cannot be used in the discourse setting and only acceptable in the literary linguistic setting.

Investigating Japanese sensation words (expressions of internal states), Kuroda (1973/ 2014) identified their unique usage in the nonreportive style. Sensation words, which express one's feelings or emotions such as happiness and sadness, are related to one's epistemology or subjectivity. In Japanese, sensation words exist in pairs, which indicate the morphological contrast between the first person and the second/third-person. For example, the word *kanashii* (sad) is morphologically classified as an adjective and only used for the first-person subject, whereas the word has to be modalized, for instance, with the suffix *garu* for the second or the third person subject. This morphological difference corresponds to the epistemological difference of the grammatical subject of a sensation word (cf. Iwasaki, 2013). Provided below are sentences, in which the word *kanashii* (sad) is utilized. These sentences

are all grammatical and can be used in the discourse setting.

(1) watashi wa kanashii
I particle sad
'I am sad.'

(2) John wa kanashi gatte iru
(name) particle sad suffix aspect:nonpast
'John seems to be sad.'

(3) Mary wa kanashii ni-chigainai
(name) particle sad certainly (logical inference)
'Mary must be sad.'

The sentences (1) and (2) respectively follow the grammatical restriction that the adjectival form *kanashii* occurs with the first-person subject and the verbal form occurs with the second/ third-person subject. The sentence (3) also follows the grammatical restriction that the adjectival form may occur with the third-person subject when the adjective form is followed by the expressions such as *ni chigainai* (must be true) or *no da* (*no*: sentence extender and *da*: copula), which indicate the existence of a subject of consciousness, who makes a judgment from outside. The sentence (3) thus is also grammatical.

However, in narratives of the nonreportive style, this grammatical restriction on the subject of a sensation word may be relaxed, which creates the unique usage of sensation words. For example, one could find a sentence in which a third-person subject occurs with the unmodalized adjective form of 'sad' as provided below.

(4) John wa kanashii
(name) particle sad
'John is sad.'

The sentence (4) is ungrammatical in the discourse setting and must be rendered, for example, as the sentence (2) or (3). However, though it is ungrammatical in the discourse setting, the sentence (4) is acceptable in narratives of the nonreportive style. This unique grammatical behavior, therefore, leads to the discussion of the interpretation of this type of sentence unique to literature.

Kuroda stated that there are two theories of how to interpret this type of unique sentence rendered in the nonreportive style: the omniscient narrator theory and the multi-consciousness theory. The omniscient narrator theory posits that the whole story is told by a single omniscient narrator as a sequence of events perceived by him. In other words, the narrator, who is identified with the speaker in discourse, is always assumed to exist throughout the text. The omniscient narrator plays a pivotal role in interpretation of the sentence (4). Taking the point of view of the subject (John), the omniscient narrator narrates John's inner states as if it were the narrator's own inner states. That is, the narrator makes an assertion about a character's internal states while the character also makes assertions about his own internal states. Kuroda, however, criticized this approach that there is no grammatical indication that the narrator exists in the text. Also, if the narrator exists throughout the text and makes an assertion about characters' internal states, the narrator uses grammar differently from the discourse setting. In other words, the language that the narrator uses is grammatically different from the language that we use. Therefore, Kuroda proposed the multi-consciousness theory, which posits that a story is comprised of structured information that is collected from different subjects of consciousness in the story including characters as well as the effaced omnipresent narrator. The narrator who has the omniscient perspective

throughout the story does not exist in this theory. In this framework, sentences like (4), Kuroda (1973/ 2014) states, is interpreted that it directly "represents the referent of the subject of the sentence (p. 49)." He explained that grammatical elements, such as morphological difference (*kanashii* and *kanashigaru*), "points semantically to the existence of a subject of consciousness whose judgment the sentence is understood to represent (Kuroda, 1973/ 2014, p55)." In other words, those grammatical elements or morphological differences have a referential force, which is directed toward a subject of consciousness and linguistically identifies the subject of consciousness. While the adjectival form *kanashii* indicates that the subject of consciousness is a first person, the modalized form *kanashigaru* indicates that due to the suffix *garu*, the subject of consciousness is a third person, who makes a judgment from outside. Kuroda states that while the omniscient narrator theory is theoretically possible and works only on the grounds that there exists the omniscient narrator, the multi-consciousness theory is linguistically supported on the grounds that a referential force of grammar, such as a morphological difference between the adjectival form *kanashii* and the adjective used with the suffix, such as *garu*, *kanashigaru*, works independently of any assumption concerning the narrator.

2.1.3. Two Frameworks for the Textual Analysis of Literature

As discussed so far, the two frameworks, which are proposed to textual analysis of literary works, are radically different in the following areas:

1) The essence of language
2) The existence of the narrator
3) Interpretation of sentences of represented internal states

These differences are all related to the notion of the omniscient narrator.

2.1.3.1. Issues concerning the Essence of Language

The framework of the communicational model, as discussed above, stipulates that the essence of language exists in communication, which entails the existence of the speaker (narrator) and the hearer (reader) in discourse as well as literature. Therefore, the notion of the omniscient narrator has been introduced. The omniscient narrator, as Pascal (1977) claims, is considered to provide the voice throughout the text. However, scholars, such as Hamburger, Banfield, and Kuroda, refuted the omniscient narrator theory for the lack of linguistic evidence as discussed earlier. As an alternative system to the communicational model, Kuroda (1975/ 2014) then proposed the multi-consciousness theory of narratives. This alternative theory of language system postulates that the essence of linguistic performance resides not in communication but in meaning-realizing acts. Kuroda explained that "a meaning-realizing act is an act of creating a meaning representation (or meaning-image, meaning) in one's consciousness (p. 90)." For example, if one hears a sequence of sound or reads a sequence of letters, the sentence exerts an effect and evokes a meaning or meaning-image in his consciousness, which is called the objective function of language in contrast with the communication function of language (Kuroda, 1975/ 2014). The objective function of language does not preclude but dominates the communicative function of language. Kuroda claimed that the speaker uses a sentence to deliver a message because he is certain that a sentence evokes a meaning in the hearer's consciousness as well as it being interpreted as communicating a message. In a discourse setting, the speaker forms and delivers a sentence to the hearer. This is where the sentence exerts the objective function on the hearer, and the sentence evokes the meaning that the speaker intends to deliver in the hearer's consciousness. In a successful communication, the meaning that the speaker intends to deliver must match the one evoked in the hearer's consciousness. In other words, in communication the objective function of language is accompanied by the communicative

function of language. However, Kuroda claimed that language does not always exert the communicative function of language because the essence of language is the objective function of language. Therefore, language can be used without a communication setting. This explains and justifies the multi-consciousness theory of narratives, which encompasses both languages in discourse and literature under one umbrella. In narratives (nonreportive/ third-person narratives), Kuroda claims, a sentence that is materialized as a sequence of letters on the page exerts the objective function of language on the reader, yet does not exert the communicative function of language. That is, the sentence creates a meaning or meaning-image in the consciousness of the reader.

2.1.3.2. Issues concerning the Existence of the Narrator

The second point that the framework of the non-communicational model challenges is the assumption of the existence of the narrator. In recent literary theories, Banfield (1982) states, the notion of the author, which was once thought of as the unity of the text, is absent from textual analysis. With the absence of the author as a concept in textual analysis, the notion of the narrator has been introduced. The narrator is a created persona in the text, which is distinct from the author. That is, the narrator is not privileged in the perspective or the voice throughout the text. Banfield states that as a consequence, the text is not unified by a single voice.

In the framework of the communicational model, the existence of the narrator is essential, which leads to the introduction of the notion of the omniscient narrator. This omniscient narrator is considered to exist throughout the text, be responsible for all the sentences in the text, and even have access to characters' internal states. This notion of the omniscient narrator, however, has been challenged for the following reasons (Banfield, 1982; Kuroda, 1975/ 2014). The first reason is the treatment of the narrator. Though the notion of the author is now absent from textual analysis, the narrator is employed as a camouflage term

for the author as the unity of the text in the framework of the communicational model. The second reason is the lack of linguistic evidence. For interpretation of sentences of represented internal states, the framework of the communicational model proposes the dual voice phenomenon, which is often presented as an evidence of the dominant narrator in the text. However, Banfield and Kuroda claimed that although the merged point of view is theoretically possible, the dual voice phenomenon does not provide any linguistic trace for the existence of the omniscient narrator in the text. In the sentence of the merged point of view, the character's point of view is grammatically embodied in the text by using linguistic elements associated with subjectivity while the narrator's point of view is not linguistically embodied in the text even though the point of view can be linguistically marked in a straightforward way. Furthermore, if the narrator makes an assertion about a character's internal state, the narrator uses grammar differently from the paradigmatic linguistic performative setting (Kuroda, 1973/ 2014). Therefore, Banfield and Kuroda state, the existence of the omniscient narrator is theoretically possible yet cannot be grammatically identified independently of assumptions that a narrator must exist.

In the framework of the non-communicational model, on the other hand, the author and the narrator are considered to be "distinct constructs of literary theory, restricting the latter to cases where the narrating poet actually does 'create' a narrator, namely the first-person narrator of the first-person narrative" (Hamburger, 1973, p. 140). In other words, the narrator exits only in the first-person narratives.[3] In this framework, the narrator, who is considered to function as one of other characters in the story, is not the locus of meaning of sentences, and subjectivity/ point of view is considered to locate in expressive grammatical

[3] Kuroda (1973/ 2014) states that there may be a case in which the narrator exists in a third-person narrative. However, in such cases, the narrator may be one of the characters in the story or effaced narrator, who can be omnipresent yet cannot be omniscient. Therefore, this narrator may not take other characters' perspectives or speak their inner states.

elements, not in the speaker. Therefore, the text of narratives is not unified by the narrator but is linguistically constituted.

2.1.3.3. Issues concerning Interpretation of Sentences of Represented Internal States

These two frameworks also govern the interpretation of sentences of represented internal states (SRIS). In the framework of the communicational model, SRIS are interpreted as merging the voice of the narrator and the voice of a character. As mentioned above, in this framework, the narrator exists throughout the text and its voice is considered to penetrate into all the sentences therein. This type of sentence also retains some features of a character's "enunciation" (Prince, 1987/ 2003). Therefore, this type of sentence simultaneously renders two different voices. However, research conducted in the framework of the non-communicational model, such as Banfield (1982) and Kuroda (1973/2014), criticized the interpretation of the dual voice phenomenon for the lack of linguistic evidence. As discussed in the sections 2.1.2.1 and 2.1.2.2, investigating sentences of represented internal states, they found a linguistic trace of a character in the text, which are linguistic elements associated with subjectivity, but they found no linguistic trace of the narrator in the text, who is identified with the speaker in discourse. In other words, this type of sentence does not linguistically indicate the existence of the speaker but directly shows the subjectivity of a character in the text. These linguistic features of subjectivity are directly represented in the text. In other words, a character's thoughts, sensation or inner states are not described by the narrator but directly shown in the text.

As discussed so far, the assumption of the existence of the narrator, which derives from the view that the essence of language exists in communication, imposes restrictions on analysis of literary texts and demands that the voice of the narrator prevails throughout a story. Under the framework of the communicational model, these conditions are automatically applied to analysis without questioning. Taking a critical view of the research

under the framework of the communicational model, the research conducted under the framework of the non-communicational model syntactically and grammatically investigated the language of literature independently of the assumption of the existence of the narrator as well as other assumptions associated with the narrator. By focusing on syntactical and grammatical behavior of the language of literature, Banfield (1982) and Kuroda (1973/ 2014, 1974/ 2014, 1975/ 2014) claimed that based on the findings of language behavior in literary texts, language in literature has its own grammar and may behave differently from language in discourse. Therefore, it should be investigated in its own right.

2.2. *The Tale of Genji* and its Textual Analysis

2.2.1. The Textual Analysis of *The Tale of Genji*: the Classic Commentary

The textual analysis of *The Tale of Genji* traces itself back to the oldest commentary known as *Genjishaku* at the end of the Heian period (the 12th century), which is the era when the tale was written. Those works written between the 12th century and the 17th century are often considered as the classic commentary, or *Kochuu* (Jinno, 2004; Mitani, 1994).[4] The classic commentary is often divided into two groups: those works written between the end of Heian period (the 12th century) and the early Muromachi period (the 14th century), and those works written between the middle of Muromachi period (the 15th century) and the early Edo period (the 17th century). Early commentaries focused on providing annotations on words, phrases, or the sources of poems for better understanding of the text. However, commentaries starting with *Kachoo Yosei*, which was written by Ichijo Kaneyoshi at the end of 15th century, shifted the focus onto the text classification of the various passages in the tale, which is considered to be one of the most significant accomplishments in the classic commentary. According to this tradition, the passages in the tale are classified into six subcategories:

[4] *Kogetsushoo*, which was written by Kitamura Kigin, is often described as the last classic commentary.

narration (*ji no bun*), the narrator's commentary (*sooshi ji*), conversation (*kaiwabun*), inner speech (*shinnai go*), letters (*shoosoku*), and poems (*waka*).[5] Narration and the narrator's commentary are attributed to the narrator while conversation, inner speech, letters, and short poems are attributed to the characters. Although narration and the narrator's commentary are attributed to the narrator, they function differently in the text. Narration moves the story forward; the narrator's commentary renders the narrator's opinions on events and characters in the tale.

It is crucial to note that these two subcategories indicate that the text classification was established on the assumption that there exists a narrator who does the telling in the tale. This assumption of the existence of the narrator stimulated further research on textual analysis of the tale, and the notion of the narrator plays a crucial role in the modern research on text analysis of *The Tale of Genji*.

2.2.2. The Textual Analysis of *The Tale of Genji*: the Late 20th Century to the Present

The current textual analysis of *The Tale of Genji* peaked in the 1970s and the 1980s. However, there was a remarkable study proposed by Tamagami in 1950-1955: *Monogatari Ondokuron*. By investigating passages classified as the narrator's commentary in the classic commentary, Tamagami claimed that the text of the tale shows the whole process of composition and reception. In the text of the tale, three strata of narration exists: first, there is a lady-in-waiting who has experienced and heard about events and transmits the tale to the next generation; second, a lady-in-waiting who writes down and edits the tale, and, third, a lady-in-waiting who narrates the tale to a princess, who looks at picture scrolls. In other words, the text of *The Tale of Genji* contains two different worlds: one world is the content of the tale (the narrated world), in which Genji and other characters exist, and the other is the

[5] The terms may vary depending on the researcher/ classic commentary.

composition and reception of the tale, in which the content of the tale is treated as a story (the narrating world). That is, the characters exist in the narrated world while these ladies-in-waiting exist in the narrating world. Investigating linguistic elements, mainly the usage of honorifics, Tamagami vigorously examined how the narrators reveal themselves in the texts.

Tamagami's theory was groundbreaking and led to the boom of the textual analysis of *The Tale of Genji* in the 1970s and the 1980s. Since the proposition of *Monogatari Ondokuron*, Mitamura (1996) states, many scholars used his analytical framework and published research on textual analysis of *The Tale of Genji*. While some studies were simple reflection and reinterpretation of Tamagami's theory, other studies adopted a critical stance toward his theory and critically questioned Tamagami's notion of the narrator (e.g., Jinno, 2004; Kurihara, 1971/ 1988; Mitani, 1989; Nakano, 1963/ 1969, 1964/ 1969; Saigo, 1983; Takahashi, 1991). They challenged Tamagami's theory that it may be not appropriate to situate a substantial narrator in the tale, which hereafter refers to a created persona in narratives. For example, the tale recounts events over the course of 70 years, and it is impossible for one narrator to witness every event. Also, in the tale, events take place in different places, such as the imperial palace and somewhere outside of the imperial palace. However, in the Heian period, people's sphere of action was limited due to the strict social norm. Moreover, there are scenes in which a character talks to herself with no one else around her, in which case it is impossible for anyone other than the reader to know what is murmured.

This question of identification of the narrator in the tale moved the textual analysis to a new level, and three main scholars, Mitani, Takahashi, and Jinno, approached this issue. Each of them proposed a new notion of the narrator, which is compatible with those issues discussed above. Where Mitani (1989) proposed "the speaker of the tale" (*washa*: 話者), Takahashi (1991) proposed "the narrator like a ghost" (*mononoke no yoona katarite*: 物の怪のような語り手), and Jinno (2004) proposed "the voice of speaking" (*wasei*: 和声). First,

Mitani's "the speaker of the tale" (*washa*) is a high-level concept of the narrator of the tale. According to Mitani, "the speaker of the tale," who is omniscient, has the penetrating perspective throughout the tale and can take all viewpoints both of the narrator and the characters. Secondly, Takahashi claimed that "the narrator like a ghost" (*mononoke no yoona katarite*) acts like a ghost in the tale. For instance, she sometimes shows herself substantially as a lady-in-waiting yet other times does not show herself clearly. In other words, the narrator like a ghost may be limited in perspective when merging into a character yet may exist as an omniscient narrator in other contexts. Lastly, Jinno's abstract concept of the narrator, namely "the voice of speaking" (*wasei*), encompasses all the ladies-in-waiting who transmit the tale, who write it down and who edit it. Introducing the notion of "the voice of speaking," he focused more on how the voices of the narrators as well as the characters are rendered in the tale. Though these scholars introduced different notions of the narrator and new frameworks of the narrative, they have not yet reached a conclusion. Also, there is no single consensus on how to cope with the two parallel worlds, one in which characters exist (the narrated world) and the other in which ladies-in-waiting do the actual storytelling (the narrating world).

In addition to the notion of the narrator, there is a crucial phenomenon that the text classification revealed: some sentences cannot be classified into any one of the subcategories proposed in the text classification. In other words, there are cases that narration or the narrator's commentary read as if they are also part of the inner speech of a character (Mitani, 1994; Jinno, 2004). This phenomenon is identified in the combinations of either narration and narrator's commentary, narration and inner speech, and narrator's commentary and inner speech. Jinno (2004) states that it is almost impossible to distinguish one category from the other in passages where this phenomenon takes place (cf. Suzuki, 1969). Sentences with this phenomenon, Mitani (1994) claims, are frequently seen in the texts of *The Tale of Genji*. This type of sentence, in fact, was already identified back in the later Edo period (the early 19th century) by Nakajima Hirotari, who referred to it as *utsuri kotoba* (うつり詞) (Jinno, 2004;

Higashihara, 1991). In the current textual analysis, this phenomenon was brought to the attention of many researchers and referred to as various names, such as *taiken wahoo* (体験話法) (Nishio, 1952/ 1991), *jiyuu-kansetsu gensetsu* (自由間接言説) or *jiyuu-chokusetsu gensetsu* (自由直接言説) (Mitani, 1994), and *naiwa-teki hyoogen* (内話的表現) (Kai ,1980). Creating a new category for the text classification, these studies tried to re-categorize the text of *The Tale of Genji*. On the other hand, instead of classification of the text, other studies such as Fujii (2001), Takahashi (1991), and Jinno (2004), focused on how the narrator and a character are amalgamated or separated in voice or perspective in this type of sentence.

A major linguistic characteristic of this type of sentence, Mitani (1994) states, is the omission of honorific expressions. Honorific expressions played a significant role in communication in the Heian period (794-1185), when the tale was written. This period was dominated by a highly aristocratic society, and it was imperative to use honorific expressions in communication. In principle, Mitani claims, honorific expressions are not used to describe one's own action or state. Therefore, passages of narration or the narrator's commentary may also be read as first-person sentences when honorific expressions are omitted for high ranked characters (Mitani, 1994, p.41). Though some studies examined other linguistic features, such as auxiliary verbs and adjectives, appearing in sentences of this phenomenon (e.g., Murakami, 1998; Stinchecum, 1980), most research on textual analysis of *The Tale of Genji* focused on the usage of honorifics (e.g., Mitani, 1989, 1994; Takahashi, 1991; Jinno 2004). Introducing the notions of free indirect discourse, Mitani claimed that this type of sentence creates two different interpretations. First, the sentence reads as if it were rendered in the first person, which makes the reader feels as if he were also experiencing the feelings or emotions of a character. Secondly, the sentence also reads as if it were rendered from the narrator's perspective. That is, free indirect discourse provides two different ways of reading. Takahashi, on the other hand, proposed the notion of psychological perspective and suggested that omission of honorifics contributes to merging the narrator and a character into one, and

inclusion of honorifics contributes to separate them. Jinno proposed the notion of dual voice and examined how different voices are layered in sentences. Although different notions of the narrator were proposed to interpret sentences containing this phenomenon, they have not come to a single consensus.

As seen in the brief overview of the previous research on textual analysis of *The Tale of Genji*, the notion of the narrator, which was established in the early commentary, has been a major focus of discussion throughout the current trend of research on the text. Though slightly different from one another in the treatment of the narrator, these studies reviewed in this section demonstrated how to treat and identify the narrator in the tale by investigating mainly honorific expressions. However, the boom in textual analysis of *The Tale of Genji* passed in the early 2000s and the field has not seen remarkable progress since.

2.3. Summary

This chapter discusses issues remaining in the textual analysis of *The Tale of Genji* as well as two different frameworks on textual analysis of the language of literature. I accomplish this by reviewing major previous studies of the tale and other research relevant to this dissertation. Through investigation of the narrator's commentary proposed in the classic commentary, the previous research on textual analysis of *The Tale of Genji* identified sentences of represented internal states appearing in the text of the tale. These previous studies were conducted under the framework of the communicational model. That is, the existence of the narrator, either substantial or insubstantial, is prerequisite for analysis. This framework, as mentioned above, dictates various conditions of analysis and interpretation of the text, which Banfield and Kuroda claim are lacking in linguistic evidence. Also, the previous studies of the textual analysis of the tale were limited mostly to the investigation of honorific expressions, which have left other linguistic elements not fully explored. This dissertation, therefore, examines the text of the tale independently of the assumption of the

prerequisite existence of the narrator with focus on linguistic elements associated with subjectivity, which, Banfield (1982) and Kuroda (1973/ 2014) claim, contribute to creating sentences of represented internal states. In order to identify linguistic elements of subjectivity to examine, the present study first conducts a quantitative analysis (Chapter 3). Based on the findings from the quantitative analysis, I perform a qualitative analysis on the text of *The Tale of Genji*, which examines sentences of represented internal states appearing in *kaimami* scenes in the tale (Chapter 4). In doing so, the present study not only examines and interprets the language of the tale based on the usage of the language in the text but also sheds new light on the textual analysis of *The Tale of Genji*.

CHAPTER 3

QUANTITATIVE ANALYSIS ON THE TALE OF GENJI

The ultimate goal of this chapter is to analyze the distribution of linguistic elements associated with subjectivity (hereafter 'subjectivity expressions') across three different text types (narration, conversation, short poem) in literary texts written in the Heian period (794-1192) and to determine lexical items associated with subjectivity to further conduct a qualitative analysis on the text of *The Tale of Genji*. In order to achieve this goal, the current chapter identifies lexical items that show a prominent appearance in each text type by examining the distribution of lexical items in a diachronic corpus established by the National Institute for Japanese Language and Linguistics, which consists of 16 literary works from the Heian period. This quantitative analysis finds three linguistic elements associated with subjectivity (auxiliary verbs, particles for exclamation and interjections) that appear prominently in both sentences of narration and conversation.

As discussed earlier, previous studies on textual analyses of *The Tale of Genji* (e.g., Mitani, 1994 and Jinno, 2004) intensively examined the usage of honorifics in sentences of narration that read as if the reader were experiencing the scene or feelings on his own (SRIS). However, in examining sentences of narration that create this phenomenon in modern English literature as well as in modern Japanese literature, different linguistic elements, which are associated with subjectivity, are found to contribute to creating this type of sentence (Banfield 1982, Kuroda 1973/ 2014). Linguistic realization of subjectivity in texts, of course, varies across languages. Due to the lack of investigation into lexical items other than honorifics in the textual analysis of *The Tale of Genji*, the current chapter analyzes mainly subjectivity expressions, specifically auxiliary verbs for speculation, particles for exclamation, and interjections. The quantitative analysis presented here was inspired by Ogiso (2015), which aims to identify word distribution across the three text types in literary works from the Heian

period. [6] The discussion in this chapter, however, goes beyond simple description of word distribution among the different text types. This chapter first establishes the criterion to determine lexical items of prominent appearance in each text type based on the frequencies of lexical items across the chronological corpus. Focusing on lexical items associated with subjectivity across the three text types, it then examines the distribution and the functions of these lexical items. Based on the distribution of subjectivity linguistic elements, the current research finally determines targeted linguistic elements associated with subjectivity and analyzes how these linguistic elements are utilized in sentences of narration, especially in sentences of represented internal states in *The Tale of Genji*. I employ the log likelihood ratio (henceforth LLR), a statistical procedure, to find the distribution of such subjective linguistics elements across different text types. As described in detail below, the LLR is suited to identify lexical items that are statistically prominent in each text type in comparison with the others; out of various types of statistical processing, the LLR is effective to compare different sized parameters.

I will first present the data and methodology that I used for the current research. In the data section, I give an overview of the Japanese language that is used in *The Tale of Genji* and explain its important linguistic aspects and present the corpus that is used in the current analysis. The methodology section consists of the LLR and the concordance search. The LLR analysis is used to identify word distribution as well as linguistic functions of lexical items in each text type (narration, conversation, and short poems). Ultimately identifying the word distribution of each text type will allow us to identify the distribution of subjective lexical items in each text type, which, Banfield (1982) and Kuroda (1973/ 2014, 1974/ 2014) claim,

[6] The CHJ of the Heian Period Series consisted of 14 literary works when Ogiso conducted his research. Also, the database was the version 1.0. The CHJ of the Heian Period Series, which the current research utilizes, has been updated to the version 2017.03. The data for the corpus is now expanded to the 16 literary works with additional two literary works, *The Gossamer Years* and *The Great Mirror*.

contribute to creating this unique phenomenon in literary texts. Based on the findings, I create concordance plots to detect where targeted linguistic features appear in the texts of *The Tale of Genji*, which leads to the next step—a qualitative analysis on the texts of the tale for close reading (Chapter 4).

3.1. Data

3.1.1. The Japanese language

Substantial written texts started to appear from the 8th century in Japan though there were some records of fragmented texts from earlier periods (Shibatani, 1990, p.119). The Japanese language is usually described chronologically in terms of five historical periods as shown in the table below.[7]

Table 3.1. The Different Stages of Japanese and Historical Period

(Reprinted from M. Shibatani, 1990, p. 119)

Old Japanese	Late Old Japanese	Middle Japanese	Early Modern Japanese	Modern Japanese
8th century	9-10-11-12 centuries	12-13-14-15-16 centuries	17-18-mid 19 centuries	mid 19-present
Nara period	Heian period	Kamakura/ Muromachi periods	Edo period	Meiji/ Taisho/ Showa/ Heisei periods

The Japanese language used in *The Tale of Genji*, which was written in the early 11th century, is classified as Late Old Japanese (794-1192) and corresponds to the political period known as the Heian period. The Heian period was crucial for the development and establishment of the Japanese writing style. Before the Heian period, Japanese did not have its own writing system and employed Chinese characters for writing starting in the fifth century (Iwasaki,

[7] There is another way to classify Modern Japanese, in which the Japanese language from middle 19th century to 1945 as Modern Japanese and the Japanese language from 1945 to present as Present-day Japanese.

2013, p.18). During the Nara and the early Heian periods, Shibatani (1990) states, two types of syllabaries were developed: *hiragana* and *katakana*. *Hiragana* was derived from the cursive form of Chinese characters; *katakana* developed from parts of these characters. The creation of these new syllabaries, *hiragana* and *katakana*, did not exclude Chinese characters from the Japanese writing system. For example, from the mid-Heian period Chinese characters were utilized in conjunction with *kana* for "mixed-Chinese character and *kana* writing" (*kanji-kana majiribun*), in which Chinese characters were "used logographically for most content words and some grammatical words" and *kana* characters, mostly *katakana*, were "used phonographically for some grammatical words and endings" (Frellesvig, 2010, p. 157). *Hiragana* was employed to write poems, calligraphy and literary works (Frellesvig, 2010; Kitahara et al., 1981). Since women, back in time, were discouraged from leaning Chinese characters, this invention of *hiragana* promoted the flourishing of women's literature in the Heian period, such as *The Tale of Genji*.[8]

. Late Old Japanese, which is distinct from Modern Japanese, had its own linguistic system. Among various differences honorifics are crucial, and honorifics were a prerequisite to successful communication. As opposed to honorifics in modern Japanese, which are used to show the speaker's respect for the person for whom the honorifics are used, honorifics in classical Japanese were used to indicate the social status of the participants in a sentence, such as the agent, patient and recipient (Nishida, 1987, p.97). There are, for example, honorifics known as absolute polite language (絶対敬語), which were only used for specific social status (Tamagami, 1959). [9] The verb *soosu* 'to say' was only used for emperors or

[8] In the literary genre known as *setsuwa*, the *kanji-kana majiribun* style was used (Frellesvig, 2010, p. 158).

[9] Kitahara (1981) states that absolute polite language proposed by Tamagami is different from the absolute polite language proposed by Kyosuke Kindaichi. Kindaichi proposes that absolute polite language refers to language use, in which honorifics are always utilized for a specific participant in a sentence regardless of it being rendered in the first, second, or third person, or in any other circumstances.

retired emperors, and the verb *keesu* 'to say' was only used for empresses or crown princes. Also, there is the usage of honorifics known as *jison-keigo* (自尊敬語), in which people of high rank, such as emperors, referred to themselves with honorifics (Nishida, 1987, 2003; Kitahara, 1981). In this case, the honorific expressions were to indicate their supreme social status. Nishida (1987) states that *jison-keigo*, which was utilized for gods or emperors, is a major characteristic of Old Japanese. However, Nishida (2003) states that starting with Late Old Japanese, *jison-keigo* was used in texts, such as literary works (e.g, *The Tale of Heike* written in the Kamakura period), diaries written by emperors, and official documents by *shoguns/ generals* though there are few instances found in *The Tale of Genji*.

Honorifics play a major role in understanding texts of classical Japanese literature (Nishida, 1987). In classical Japanese, participants in a sentence, such as the agent, patient, and recipient, are often absent, and characters are rarely mentioned by their names, especially when they are high-ranked characters. Honorifics are crucial in order to identify participants in a sentence as well as the speaker who produces an utterance. In principle, forms of respect indicate that the agent in a sentence is high-ranking and forms of humbleness indicate that the recipient or patient in a sentence is high-ranking. [10] Therefore, analyzing the use of honorifics is crucial in understanding the texts of classical Japanese.

3.1.2. The Corpus of the Japanese Language

The current research utilizes part of the Corpus of Historical Japanese (henceforth the CHJ), which is a diachronic corpus of Japanese language provided by the National Institute for Japanese Language and Linguistics (henceforth NINJAL). NINJAL launched the

[10] Nishida claims that honorifics used in conversations or letters do not necessarily follow the social hierarchy. For instance, honorifics were used for a lower status to show intimacy or affection in conversations or letters. In principle, honorifics used in narration, on the other hand, follow the social hierarchy; honorific expressions should be used to describe high-ranked characters.

KOTONOHA project to create a large database of Japanese language based on the need for language research. The KOTONOHA project aims to encompass not only Present-day Japanese, both spoken and written, but also samples from Old Japanese (710-794) to Modern Japanese (1868-1945). Since it is impossible to develop a single corpus that consists of all the Japanese languages from different periods as well as written and spoken forms, several sub-corpora have been developed. The CHJ is one such corpus, which was made available to the researchers in 2009 and is still expanding.[11] The CHJ is annotated with pronunciation, categorization of parts of speech (POS), and morphological information (cf. Ogiso, 2015), and currently consists of materials from the Nara period (710-794) through the Taisho period (1912-1926). The CHJ is accessible through the online search engine *Chunagon* provided by NINJAL.[12]

For the current research, I utilize the CHJ of the Heian Period Series, a sub-corpus of the CHJ, which covers literary pieces written in Late Old Japanese (794-1192). The CHJ of the Heian Period Series, the texts of which are based on the *Shinpen Nihon koten bungaku zenshū*, consists of various literary works from diverse genres, such as anthologies of poems, diaries, and narratives. The corpus currently consists of the following 16 *kana* literary works:

1. *The Tale of the Bamboo Cutter* (*Taketori Monogatari*)
2. *The Collection of Japanese Poems of Ancient and Modern Times* (*Kokinwakashu*)
3. *The Tales of Ise* (*Ise Monogatari*)
4. *The Tosa Diary* (*Tosa Nikki*)
5. *The Tales of Yamato* (*Yamato Monogatari*)
6. *The Tales of Heichu* (*Heichu Monogatari*)

[11] In order to utilize the CHJ, subscription is required. For more information, refer to the official NINJAL website: https://pj.ninjal.ac.jp/corpus_center/chj/subscription-en.html

[12] The maximum limit of lines downloadable by the *Chunagon* corpus search engine is one million lines.

7. *The Tale of Ochikubo* (*Ochikubo Monogatari*)
8. *The Pillow Book* (*Makura no Soshi*)
9. *The Tale of Genji* (*Genji Monogatari*)
10. *The Diary of Lady Murasaki* (*Murasaki Shikibu Nikki*)
11. *The Sarashina Diary* (*Sarashina Nikki*)
12. *The Tales of the Riverside Middle Counselor* (*Tsutsumi Chunagon Monogatari*)
13. *The Diary of Sanuki no Suke* (*Sanuki no Suke Nikki*)
14. *The Gossamer Years* (*Kagero Nikki*)
15. *The Diary of Izumi Shikibu* (*Izumi Shikibu Nikki*)
16. *The Great Mirror* (*Okagami*)

Besides pronunciation, parts-of-speech (POS), and morphological information, the CHJ of the Heian Period Series is also annotated with style (narration, conversation, and short poems). In addition, the speaker of each dialogue is also annotated in the texts of *The Tale of Genji*.

Provided below is the description of the CHJ of the Heian Period Series for the 16 literary works.[13]

[13] National Institute for Japanese Language and Linguistics (2017) "Corpus of Historical Japanese, Heian Period Series." (Short Unit Word data 1.1 / Long Unit Word data 1.1)
http://pj.ninjal.ac.jp/corpus_center/chj/heian.html

Table 3.2. Description of the CHJ of the Heian Period Series (2017.3. version)

	Narration	Conversation	Short Poems	Classic commentary	Captions	Letter	Total
The Tale of the Bamboo Cutter	5224	4463	255	0	0	374	10316
The Collection of Japanese Poems of Ancient and Modern Times	2461	0	18628	873	9298	0	31260
The Tales of Ise	9200	536	4048	0	0	41	13825
The Tosa Diary	5077	549	1059	0	0	0	6685
The Tales of Yamato	15586	2496	4979	0	0	30	23091
The Tales of Heichu	7724	1864	2815	0	0	0	12403
The Tale of Ochikubo	27891	25515	1180	0	0	0	54586
The Pillow Book	51728	13707	602	0	0	0	66037
The Diary of Izumi Shikibu	5124	3302	2465	0	0	0	10891
The Tale of Genji	297632	132597	13123	0	0	2359	445711
The Diary of Lady Murasaki	16160	989	293	0	0	0	17442
The Tales of the Riverside Middle Counselor	9508	5375	813	0	0	0	15696
The Sarashina Diary	9964	3217	1479	0	0	0	14660
The Diary of Sanuki no Suke	11386	3789	369	0	0	0	15544
The Great Mirror	1611	67807	1849	0	0	0	71267
The Gossamer Years	31009	10080	6175	0	0	0	47264
Total	507285	276286	60132	873	9298	2804	856678

Table 3.2. shows the numbers of word tokens for each literary work. The texts of each literary work are categorized in six styles: narration, conversation, short poems, classic commentary, captions, and letters.[14] As seen above, most texts are categorized in narration, conversation, or short poems. These categorizations, Ogiso (2015) states, are all made based on the annotation/ information of the texts of the *Shinpen Nihon koten bungaku zenshū* (p.94). Short poems consist of 31 syllables – divided line by line in the pattern of 5-7-5-7-7. While texts that are bracketed in the texts of the *Shinpen Nihon koten bungaku zenshū* are classified as conversation, texts that are not included in any of the five categories (conversation, short poems, classic commentary, captions, and letters) are classified as narration. In this corpus, however, there are no such categories for inner speech or the narrator's commentary that are proposed in the studies of *The Tale of Genji*. Since the categories, classic commentary, captions, and letters, appear only in specific literary works, the current study only examines for the quantitative analysis the texts that are categorized in narration, conversation, and short

[14] Classic commentary, also known as *kochuu* (古注) in Japanese, are commentaries that were written in the past. Captions are texts usually placed before a short poem that explain such things as circumstances in which a short poem was composed.

poems.[15]

3.2. Methodology

The current research adopts the methodology used in corpus linguistics, which has developed over the past few decades for linguistic analyses. Corpus linguistics is an empirical approach to linguistic research and provides evidence for the systematicity of language variation, which can be found by utilizing the empirical methods. Corpus linguistic research analyzes different aspects of language variation as well as language use, such as the usage of various parts of speech (e.g., the use of synonyms) and the usage of various grammatical structures (e.g., the use of active voice and passive voice). Also, corpus linguistic research uses large amounts of textual data as the basis of analysis. It also extensively utilizes the computer to analyze large amounts of textual data. These characteristics of corpus research provide a benefit. Using the computer enables us to analyze large data that would otherwise be unfeasible. By doing so, findings that are obtained from analyzing a large quantity of data provide more reliable generalizability and validity. Corpus linguistic research, Biber (2010) states, often yields unexpected linguistic patterns, which are counter to a native speaker's intuition for language use, or discovers linguistic constructs unidentified by traditional linguistic theory.

There are two different approaches in corpus linguistics: the corpus-based approach and the corpus-driven approach. Studies using the corpus-based approach aim to describe and explain linguistic patterns of variation and usage. That is, a major goal of the corpus-based approach is to discover the systematic linguistic patterns that pertain to the linguistic features identified in standard linguistic theory. For example, Biber et al. (2016) identify different ways of elaboration in conversation as well as in academic writing by investigating the

[15] There is no category for inner speech or the narrator's commentary, which are proposed in the studies of *The Tale of Genji*. These categories may be included in texts of narration.

distribution of complement clauses, such as *that*-clauses and *WH*-clauses, and phrasal modifiers, such as attributed adjectives and prepositional phrases. In order to elaborate sentences, complement clauses are preferred in conversation while phrasal modifiers are preferred in academic writing. That is, sentences in conversation tend to be structurally complex, whereas sentences in academic writing tend to be structurally compressed. On the other hand, the corpus-driven approach is more inductive and utilizes a corpus to identify linguistic phenomena that have not been previously recognized. Tognini-Benelli (2001) claims that the "descriptions aim to be comprehensive with respect to corpus evidence" (p. 84). In other words, in studies using the corpus-driven approach the corpus functions as an empirical basis from which the data is systematically derived, without any prior assumptions and expectations, "from the recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Benelli 2001, p. 87). Therefore, conclusions or claims are made exclusively based on the observation of data extracted from the texts. Biber (2009), for instance, examines linguistic patterns of multi-word sequences in conversation as well as academic writing. Investigating most frequent multi-word sequences in both categories, he finds that conversation indicates a preference for "fixed continuous sequences of words" while academic writing shows a preference for "formulaic sequences of words with an internal variable slot."

Research of corpus linguistics was restricted due to the lack of computational technology. However, it is now possible to process a large corpus of texts by using the computer. The use of a corpus for literary text analysis provides the following benefits. First, the use of a corpus accounts for the texts from an empirical perspective, which is derived from a substantial amount of data. Secondly, corpus-driven research can integrate both quantitative research and qualitative research. For instance, the data derived from the corpus may be further investigated from the perspective of theoretical linguistic analysis, and the findings may be further examined closely in actual literary texts. Lastly, replication research

may be conducted in order to scrutinize the results of the cornerstone research, which may verify the validity of research.

The two corpus linguistic approaches mentioned above (the corpus-based approach and the corpus-driven approach) can also be applied to investigation of literary texts. The corpus-based approach may be useful to investigate collocation of words in short poems. The corpus-driven approach, on the other hand, is effective for the current research, which aims to identify linguistic elements that create sentences that read as if the reader were experiencing the scene or feelings on his own. The previous research on textual analyses of *The Tale of Genji*, as discussed before, has been limited to qualitative analysis using the methodology of close reading of literary texts to investigate the textualization of the narrative. Two recent innovations, however, provide an opportunity to make a shift in this research tradition. One is the new development in corpus linguistics and the other is availability of the corpus of classical Japanese literature that comes from the same time period of *The Tale of Genji*. The corpus approach is especially suited to study the language and discourse of Late Old Japanese as it is impossible to collect new data more than already existing data or to conduct experiments to confirm findings. Using large data enables us to take an inductive approach to the data and to discover linguistic phenomena that have not been discussed in traditional linguistic studies. The quantitative analysis thus will provide a new perspective and shed light on the text analysis of the tale.

3.3. Quantitative Analysis 1: Log Likelihood Ratio (LLR)

In this section, I conduct a statistical analysis on 16 literary works from the Heian period. In doing so, I attempt to identify prominently appearing linguistic elements by examining the word distribution across the three text types (narration, conversation, short poems). This statistical analysis particularly focuses on lexical items associated with subjectivity, which is a crucial linguistic function in creating sentences of represented internal

states (Banfield, 1982; Kuroda, 1973/ 2014). Statistical processing plays a significant role in this quantitative analysis. Comparison of the actual numbers of appearance of a lexical item in each text type does not show how much more or less frequently the lexical item appears in one text type over the other, because the word tokens (the size of the population parameters) are different in each text type. A simple solution to this difficulty is conversion of the raw token frequencies into the frequencies per million. This solution has two advantages. First, it is possible to put lexical items in order based on the frequencies per million; secondly, it is also possible to compare the frequencies of a lexical item in different text types. However, this solution also has two disadvantages. The first disadvantage is that the ranking of the frequencies per million does not suggest how more or less frequently the lexical item appears in one text type over the other. The second disadvantage is that high-frequency lexical items become predominant in ranking. That is, low-frequency lexical items are automatically ranked low. An alternative approach to overcome these problems is to use a statistical method. Though there are various statistical methods available, such as the chi-square, LLR works more effectively for the current study. Dunning (1993) claims that LLR works effectively in such cases where different sized corpora, meaning different sized population parameters, are compared or where the expected frequency becomes low or medium.[16] The LLR value indicates how more or less frequently a given lexical item is used in one text type (the targeted text type) compared to the predicted frequency of the lexical item in the other text type (the reference text type).[17] In other words, if the LLR value of a given lexical item is positive and is also high, the lexical item appears more frequently in the targeted text type compared to the expected frequency of the targeted lexical item in the reference text type. On

[16] Dunning states that the chi-square value becomes unreliable in such cases.

[17] For example, when we calculate LLR values of the lexical items in the text type of narration, the targeted text type is the text type of narration and the reference text types are the text types of conversations and short poems.

the other hand, if the LLR value of a given lexical item is negative as well as low, the lexical item appears less frequently in the targeted texts compared to the predicted frequency of the lexical item in the reference text type. Provided below is the formula for LLR.[18]

The variable a indicates the frequency of the word in the targeted text type.
The variable b indicates the frequency of the word in the reference text type.
The variable c indicates the total number of the words in the targeted text type minus a.
The variable d indicates the total number of the words in the reference text type minus b.

$$LLR = 2[a \text{ Log } a + b \text{ Log } b + c \text{ Log } c + d \text{ Log } d - (a+b)\text{Log}(a+b) - (a+c)\text{Log}(a+c) \\ -(b+d)\text{Log}(b+d) - (c+d)\text{Log}(c+d) + (a+b+c+d)\text{Log}(a+b+c+d)]$$

In the above equation, the term $a \text{ Log } a$ should be replaced with 0 when $a = 0$ and $b \text{ Log } b$ with 0 when $b = 0$, and the right-hand side should be multiplied by -1 when $\frac{a}{a+c} < \frac{b}{b+d}$.

3.3.1. Log Likelihood Ratio on the Texts of the 16 Literary Works

In this section, I conducted an LLR analysis to identify lexical items that show a prominent appearance in the three different text types (narration, conversation, and short poem). In other words, LLR analysis determines lexical items that appear more or less frequently in one text type over the other text types. The analysis is conducted on the 16 literary works in the CHJ of the Heian Period Series; first on the 15 literary works except for *The Tale of Genji,* and then *The Tale of Genji* itself separately. The LLR analysis on the 16 literary works reveals the overall distribution of lexical items and finds lexical items of prominent appearance in each text type in the corpus. However, almost 50% of the data of CHJ of the Heian Period Series consists of *The Tale of Genji*, so it is necessary to investigate whether the results of the 16 literary works represent overall linguistic tendencies. Therefore, the current section compares the results of the 15 literary works and *The Tale of Genji* to

[18] See Kilgariff (2001) for more details.

those of LLR analysis on the 16 literary works.

Provided below is the description of the CHJ of the Heian Period Series for the 16 literary works. Table 3.3. below provides the detailed information of narration, conversation, and short poems.[19]

Table 3.3. Description of the CHJ of the Heian Period Series: Three Text Types

literary works	narration	conversation	short poem	
The Tale of the Bamboo Cutter	5224	4463	255	
The Collection of Japanese Poems of Ancient and Modern Times	2461	0	18628	
The Tales of Ise	9200	536	4048	
The Tosa Diary	5077	549	1059	
The Tales of Yamato	15586	2496	4979	
The Tales of Heichu	7724	1864	2815	
The Tale of Ochikubo	27891	25515	1180	
The Pillow Book	51728	13707	602	
The Diary of Izumi Shikibu	5124	3302	2465	
The Tale of Genji	297632	132597	13123	
The Diary of Lady Murasaki	16160	989	293	
The Tales of the Riverside Middle Counselor	9508	5375	813	
The Sarashina Diary	9964	3217	1479	
The Diary of Sanuki no Suke	11386	3789	369	
The Great Mirror	1611	67807	1849	
The Gossamer Years	31009	10080	6175	
Total	507285	276286	60132	
The total number of words				843703

The CHJ of the Heian Period Series for the 16 literary works consists of 843,703 word tokens and 11,423 lemmas[20]. As seen above, the size of the population parameter varies across the three text types. Narration consists of 507,284 lexical items, while conversation consists of 276,286 lexical items and short poem consists of 60,133 lexical items.

First, I arrange the words of the CHJ of the Heian Period Series into high frequency order for each text type (narration, conversation, and short poem). I then process each list

[19] Refer to Table 3.2. in Chapter 3 for more details of the CHJ of the Heian Period Series.

[20] A lemma is a representative form of a lexical item that is usually used as the headword in a dictionary.

statistically based on LLR, which determines the LLR value of each lexical item in each text type. Figure 3.1 provided below shows the distribution of the LLR values for narration of the 16 literary works. The vertical axis indicates the numbers of lexical items while the horizontal axis indicates the values of LLR. A value of LLR is not impacted by its actual number of occurrences whether it is small or large in number. If a lexical item has an LLR value that is at 0.0 or close to 0.0, the lexical item is not statistically prominent. In other words, it is almost equally utilized across all the text types. On the other hand, if a lexical item has an LLR value that is high, either positive or negative, it appears either more frequently or less frequently in one text type over the others.

Figure 3.1. Distribution of the LLR Values for Narration of 16 literary works

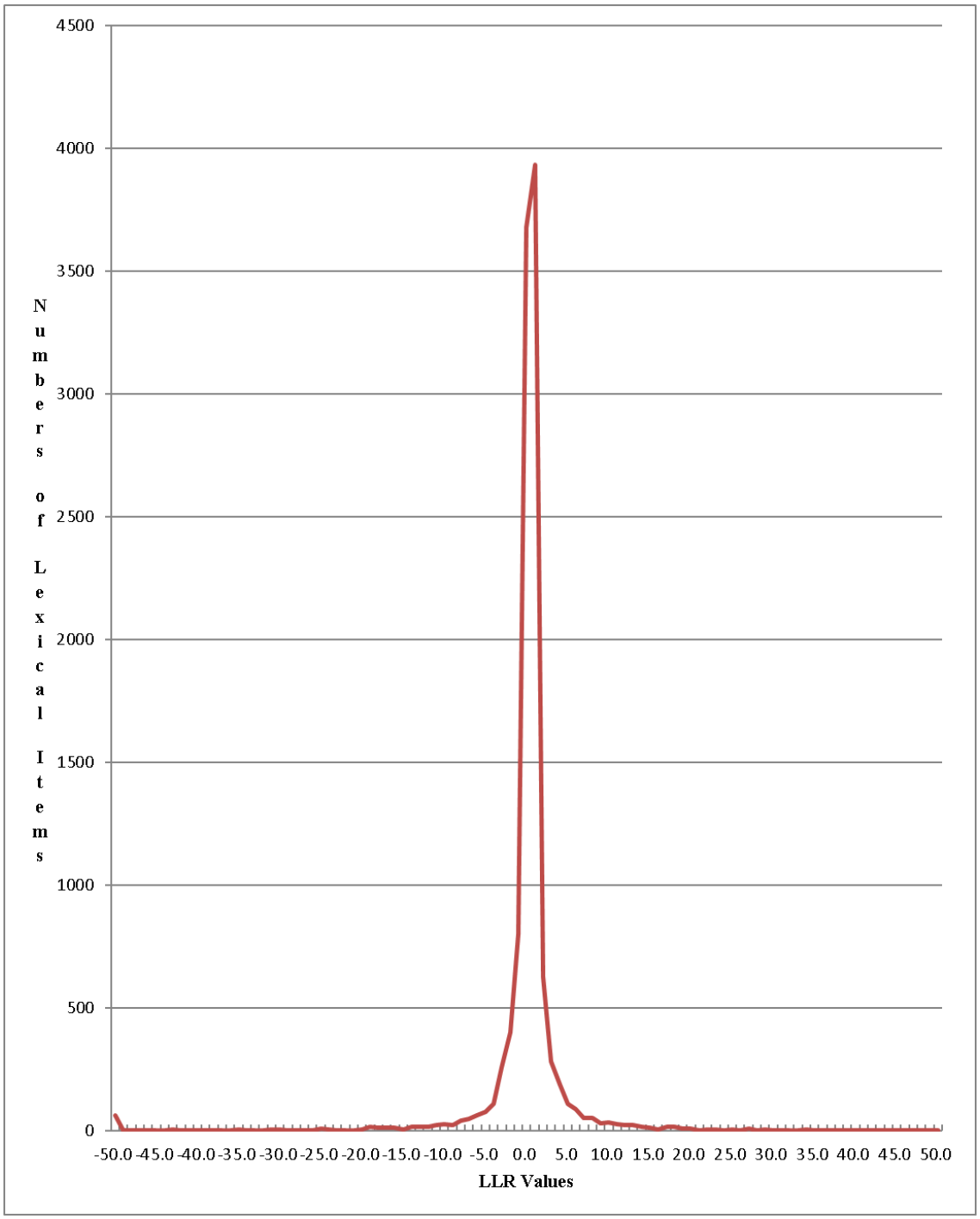
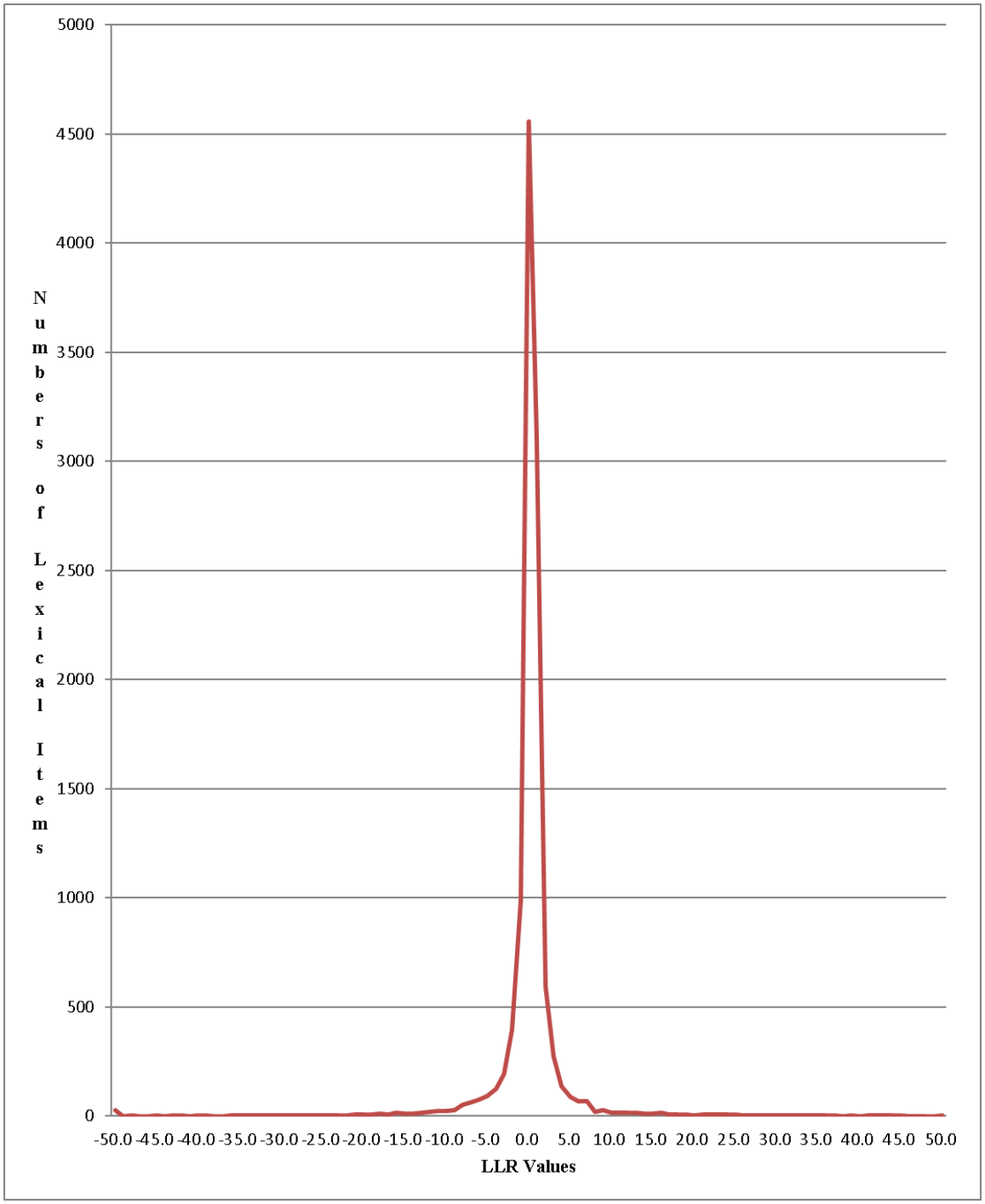


Figure 3.1 above shows that a large number of lexical items in the text type of narration have LLR values that are close to 0.0, which suggests that most lexical items are not statistically prominent in this text type. Next, we examine the results of the text types of conversation and short poem. Figure 3.2 provided below shows the distribution of the LLR values for

conversation of the 16 literary works.

Figure 3.2. Distribution of the LLR Values for Conversation of 16 literary works

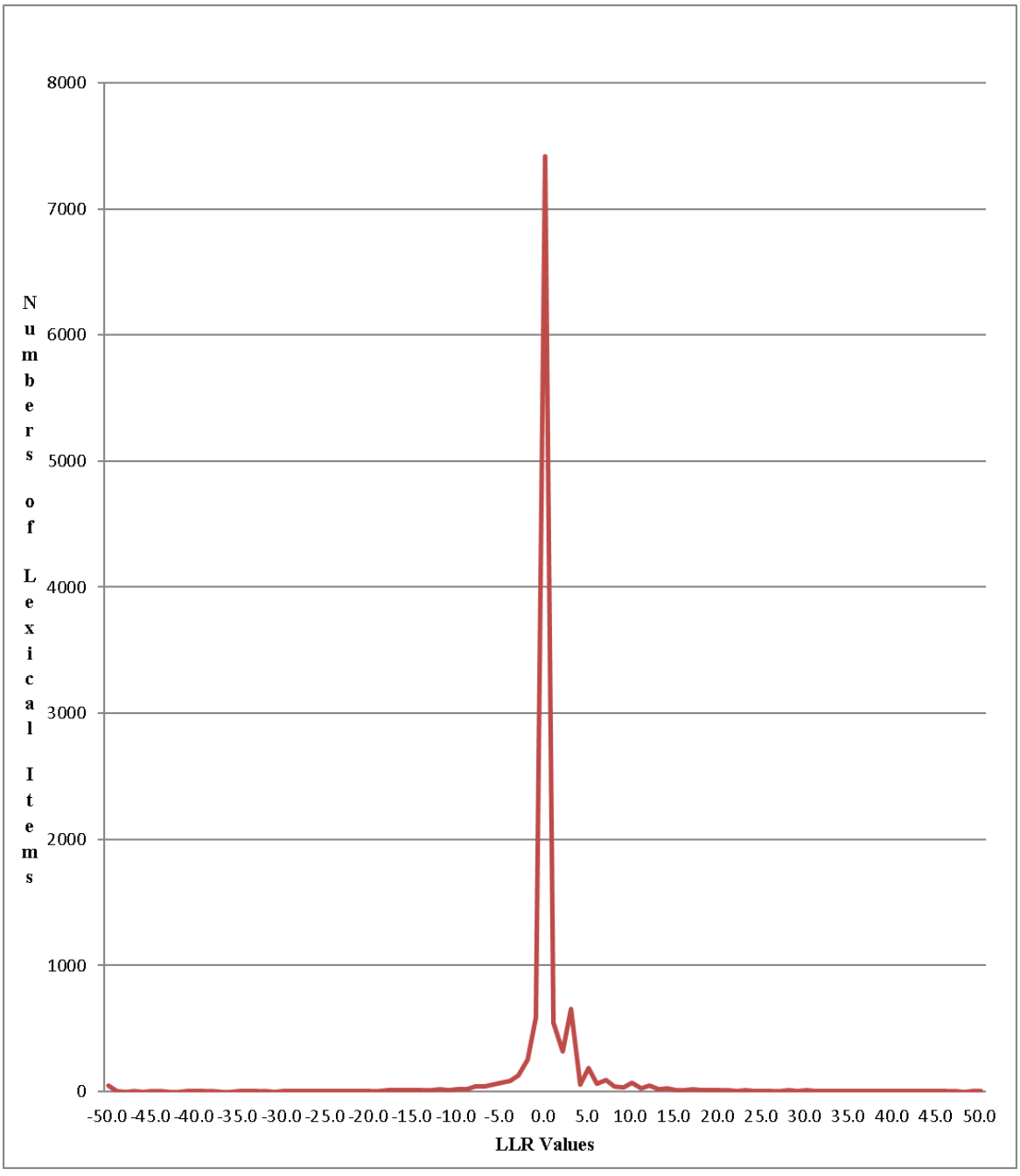


The result of the LLR distribution in conversation is similar to that of narration, which is depicted in Figure 3.1. Most lexical items have LLR values that are at 0.0 or close to 0.0. Yet,

some lexical items have high values of LLR, either positive or negative, which suggests that these lexical items appear more frequently or less frequently in conversation over the others.

Figure 3.3. provided below shows the distribution of the LLR values for short poems of the 16 literary works.

Figure 3.3. Distribution of the LLR Values for Short Poem of 16 literary works



Finally, the result of the LLR distribution in short poems is presented above. Again, most lexical items have LLR values that are at 0.0 or close to 0.0 while some lexical items have high LLR values that are either positive or negative. That is, most lexical items do not show a preference with which they appear in particular text types, whereas some lexical items appear more prominently in the text type of short poem over the others.

To provide a brief summary of the three figures presented above, a large number of lexical items have LLR values that are very close to 0.0 across the three text types, which suggests that most lexical items do not have a preference. In other words, these lexical items are almost equally utilized across all the text types. However, there are also lexical items that are located widely spread out across different LLR values though they are small in number. These lexical items that have high positive or negative values of LLR, as explained earlier, are the ones that appear more frequently or less frequently in one text type over the others. In other words, these lexical items of high LLR values show dominant or significant appearance in one text type over the others. To decide from which LLR value lexical items show prominent appearance in a given text type, a cut-off point has to be made. In order to determine a cut-off point, I employ standard deviation.[21] Standard deviation, which is a means to measure how spread out the data is, allows for determining a cut-off point based on a single criterion considering the word distribution of each text type. One standard deviation indicates the average distance from the mean value of the LLR of all the lexical items in a given textual type. The values of lexical items within one standard deviation (between the mean value plus the value of the standard deviation and the mean value minus the value of the standard deviation) are considered to be not either exceptionally high or exceptionally low; the values of lexical items outside of one standard deviation are considered to be

[21] The p-value (probability value) to reject the null hypothesis is extremely small compared to the commonly used significance level threshold (0.01-0.05). In other words, it is not practical to apply commonly used statistical tests to evaluate a cut-off point for the current study.

prominent. In other words, those lexical items that are included within one standard deviation are defined to appear in all the text types more or less equally; they do not show significant appearance in a particular text type. Those lexical items that are located outside of one standard deviation tend to appear more frequently in one text type over the others (Highest Frequency) or tend to appear less frequently in one text type over the others (Lowest Frequency). In other words, they show prominent appearance in a specific text type.

3.3.2. Distribution of Lexical Items among Different Text Types

I now investigate lexical items that are located outside of one standard deviation in the three text types (narration, conversation, and short poem). In the appendix are provided the full lists of those lexical items in each text types, including the lexical item, the actual number of appearances of each lexical item in each text type, the rank, parts of speech (a subcategory is provided for particles,)[22] and the LLR values.

For narration, the mean value is -0.559 and the standard deviation is 30.567; 98.7% of the lexical items of narration are included within the standard deviation. That is, 1.3% of the lexical items of narration show prominent appearance in narration. 57 lexical items are classified as highest frequency words for narration (henceforth HFq (narration), i.e., these are preferred words for narration. In contrast, 91 lexical items are classified as lowest frequency words for narration (henceforth LFq (narration)), i.e., these are dispreferred words for narration. Provided below is the list for preferred words for narration.

[22] The subcategory is equivalent to *Chuubunrui* in the CHJ of the Heian Period Series.

Table 3.4. List of Preferred (highest frequency) Words for Narration

parts of speech	frequency
adjective	4
particle	8
auxiliary verb	3
prefix	2
suffix	6
verb	17
adverb	1
noun	16
Total	**57**

Most lexical items classified as HFq (narration), as Ijima (2011) states, do not have the communicative function (e.g., request and prohibition) or the expressive function (e.g., lexical items for exclamatory) because these linguistic elements require a communication setting, which requires the speaker and the addressee. That is, the function of these linguistic elements is not suitable for the function of the text type of narration, which is "moving a story forward" (Hopper and Thompson, 1980). Most lexical items classified as HFq (narration), therefore, are dedicated to describing characters or producing continuous sentences. There are lexical items listed across different parts of speech that are utilized to describe an object from outside: the suffixes *ge* ‘to appear to’ and *garu* ‘to show signs of’ and the nouns *kehai* ‘appearance,' *sama* ‘appearance,' and *keshiki* ‘appearance.' Out of eight particles, five particles (*te*, *ni*, *ba*, *tsutsu*, and *do*) are classified as conjunction, which contribute to producing continuous sentences in narration; three particles (*to*, *tote* and *nado*) are utilized as citational markers to cite characters’ conversations or inner speeches. As for auxiliary verbs, the auxiliary verbs *ri* and *tari* indicate resultative, continuative or perfective, while the auxiliary verb *keri* indicates hearsay past or exclamatory recognition. Most verbs describe specific actions to describe characters’ actions, such as *kaku* ‘to write,' *naku* ‘to cry,' and *warau* ‘to laugh.' Four verbs are classified as honorifics: *tamau* ‘to indicate respect,' *obosu* ‘to

think,' *notamau* ‘to say,' and *kikoyu* ‘to say.' The verb *tamau*, *obosu*, and *notamau* are classified as forms of respect whereas the verb *kikoyu* is classified as a form of humbleness. Also, the adjective *okashi* ‘to be intriguing,' which is a representative word in the Heian literature, is listed in HFq (narration).

Provided below is the list of lowest frequency words for narration.

Table 3.5. List of Dispreferred (lowest frequency) Words for Narration

parts of speech	frequency
interjection	3
adjective	2
particles	18
auxiliary verb	12
suffix	3
pronoun	9
verb	14
adverb	9
noun	21
Total	**91**

Lexical items classified as LFq (narration), as opposed to the lexical items listed in HFq (narration), have the communicative function or the expressive function. Three interjections are listed: *ana*, *ide*, and *iza*, none of which are listed in HFq (narration). Out of 18 particles, the particles *tomo*, *o*, and *domo* are classified as conjunction, yet the other particles have the expressive function or the communicative function. For example, the particles *kana* and *yo* are used for exclamatory while the particle *ya* and *ka* are used for question. As for auxiliary verbs, seven out of 12 auxiliary verbs are categorized as modals: *mu*, *ji*, *nari*, *ramu*, *beshi*, *mashi*, and *meri*. The other auxiliary verbs have different functions. The auxiliary verb *ki* indicates past and the auxiliary verbs *sasu* and *su* indicate causative or honorific; the auxiliary verbs *tsu* and *nu* indicate perfective. Among 14 verbs, seven verbs are classified as honorifics: *haberi* ‘to show politeness,' *tamau* ‘to indicate humbleness,' *owashimasu* ‘to

exist,' *moosu* ‘to say,' *uketamawaru* ‘to receive,' *makaru* ‘to go,' and *oboshimesu* ‘to think.' Two verbs are classified as forms of respect: *owashimasu* and *oboshimesu*, while the other five verbs are classified as either forms of humbleness, such as *tamau* and *moosu*, or forms of politeness, such as *haberi*. The honorific *haberi*, which has the communicative function, shows respect to the addressee. Also, there are nouns for deixis, *ima* ‘now,' *koko* ‘here,' and *kore* ‘this.'

Distribution of words in the text type of conversation shows a sharp contrast against those in the text type of narration. For conversation, the mean value is 0.547 and the standard deviation is 33.077. 99.0% of the lexical items are included within the standard deviation. That is, only 1% of the lexical items of conversation show significant appearance in conversation. 76 lexical items are classified as highest frequency words for conversation (henceforth HFq (conversation)). Table 3.6. shows distribution of these words across different parts-of-speech. Provided below is the list of highest frequency words for conversation.

Table 3.6. List of Preferred Words for Conversation

parts of speech	**frequency**
interjection	**2**
adjective	2
particle	11
auxiliary verb	10
suffix	2
pronoun	10
adverb	10
noun	17
verb	12
Total	**76**

Lexical items of HFq (conversation) have the communicative function or the expressive function, which is suitable for the text type of conversation. There are two interjections listed,

ana and *ide*. Not surprisingly, both of which are listed in LFq (narration). In other words, what is frequent in the text type of conversation is not frequent in the text type of narration. Out of 11 particles, the particle *tomo* is classified as conjunction, while the other particles have either the expressive function or the communicative function. For example, the particle *na* indicates prohibitive in conjunction with the particle *so* whereas the particles *ya*, *yo*, and *kana* indicate exclamatory. As for auxiliary verbs, six out of ten are modals: *mu*, *beshi*, *nari*, *meri*, *ji*, and *maji*. Among the other auxiliary verbs, *ki* indicates past and *su* and *sasu* indicate causative or honorific while *nu* indicates perfective. Out of 12 verbs, nine verbs are classified as honorifics: *haberi* 'to show politeness', *tamau* 'to indicate humbleness', *owashimasu* 'to exist', *moosu* 'to say', *uketamawaru* 'to receive', *makaru* 'to go', *saburau* 'to serve', *oboshimesu* 'to think', and *tatematsuru* 'to give'. [23] The honorific verb *haberi*, as mentioned above, is used to express politeness to the addressee. Except for the verbs *owashimasu* and *oboshimesu*, the verbs classified as honorifics are either forms of humbleness or forms of politeness. Also, there are pronouns and nouns that indicate deixis: *ko* 'this,' *koko* 'here,' *kore* 'this,' and *ima* 'now.'

42 lexical items are classified as lowest frequency words for conversation (henceforth LFq (conversation)). Table 3.7 provided below shows the distribution of lexical items of LFq (conversation) across different parts-of-speech.

[23] There are two different verbs of *tamau*: one is a verb of form of respect and the other is a verb of form of humbleness. They conjugate differently.

Table 3.7. List of Dispreferred Words for Conversation

parts of speech	frequency
adjective	1
particle	7
auxiliary verb	3
prefix	1
suffix	2
verb	19
noun	9
Total	**42**

Lexical items classified as LFq (conversation), contrary to HFq (conversation), do not have the communicative function or the expressive function. In other words, these linguistic elements do not require the speaker and addressee setting. Among seven particles, four particles (*te*, *ba*, *tsutsu*, and *ni*) are classified as conjunctions; two particles (*to*, *tote* and *nado*) are utilized as citational markers. As for auxiliary verbs, *ri* and *tari* indicate resultative, continuative or perfective whereas *keri* indicates hearsay past or exclamatory recognition. Two verbs classified as honorifics are *obosu* 'to think' and *notamau* 'to say,' both of which are forms of respect. The other verbs describe specific actions, such as *iu* 'to say,' *naku* 'to cry' and *kaku* 'to write.' Also, the only adjective listed is *okashi*, which is a representative word in the Heian literature.

We now move on to examine lexical items that appear in the text type of short poem. For short poems, the mean value is 0.530 and the standard deviation is 20.938. 97.2% of the lexical items are included within the standard deviation. That is, 2.8% of the lexical items of short poems show prominent appearance in short poems. Among these lexical items, 211 lexical items are classified as highest frequency words for short poems (henceforth HFq (short poem)). Provided below is the list of highest frequency words across different parts-of-speech for short poem.

Table 3.8. List of Preferred Words for Short Poem

parts of speech	frequency
adjective	5
particle	17
auxiliary verb	8
suffix	2
pronoun	4
verb	49
noun	126
Total	**211**

Compared to the two text types seen above (narration and conversation), more lexical items are classified as HFq (short poem). The increase of lexical items for HFq (short poem) is due to diction of short poems. The words utilized in poetry are known as *uta kotoba* (poetic words), which are different from everyday words. In poetry, it is considered improper to use unrefined words or direct expressions. For instance, when depicting feelings or emotions, it is considered best to use nature metaphors instead of describing the emotions directly. Therefore, many lexical items listed as HFq (short poem), especially nouns, are related to nature or classified as poetry words. Among them, some lexical items are usually used for pivot words (*kakekotoba*), which is one of the common techniques utilized in short poems, where one word represents two homophonic words. For instance, *matsu* can be either the noun *matsu* ‘pine tree’ the verb *matsu* ‘to wait.' Also, some lexical items have the expressive function for delivering the composer’s feeling. The nouns are listed more frequently than other parts of speech among lexical items in the HFq (short poem). Most nouns may be categorized into nature, plants, clothes, and animals. For example, *shigure* ‘shower,' *tsuyu* ‘dew,' and *kawa* ‘river’ are lexical items for nature; *sode* ‘sleeves,' *koromo* ‘robes,' and *tamoto* ‘sleeves’ are lexical items relating to clothes. Other nouns are associated with animals, such as *hototogisu* ‘a cuckoo,' *uguisu* ‘nightingale,' and *shika* ‘deer’ while others are associated

with plants, such as *matsu* ‘pine trees,' *ominaeshi* ‘parthenia flowers,' and *ume* ‘Japanese apricot/plum.' These nouns listed are all poetry words.[24] Verbs do not seem to have biased semantic features as all the verbs describe specific action, such as *au* ‘to meet,' *matsu* ‘to wait,' and *saku* ‘to bloom.' However, as explained before, some of them are commonly used as *makurakotoba* ‘pivot words.’ As for auxiliary verbs, five out of eight are classified as modals: *ramu* (present speculation; speculation about a cause), *rashi* (evidential supposition), *beranari* (speculation), *ji* (negative speculation), and *mashi* (counterfactual speculation; desire for a hypothetical state). There are other auxiliary verbs with different functions. The auxiliary verb *zu* indicates negative and the auxiliary verb *nu* indicates perfective while *ki* indicates past. As for particles, some particles are classified as conjunctions, such as *domo*, *tomo*, and *o*, and other particles have the expressive function, such as *ya* and *kana* for exclamatory.

114 lexical items are classified as lowest frequency words for short poems (henceforth LFq (short poem)). Table 3.9 below shows the distribution of the lexical items as LFq (short poem) across different parts-of-speech.

[24] These nouns are all listed as poetry words in Katagiri (1999).

Table 3.9. List of Dispreferred Words for Short Poem

parts of speech	frequency
adjective	10
particle	7
auxiliary verb	8
conjunction	2
prefix	6
suffix	6
pronoun	4
verb	27
adverb	10
noun	34
Total	**114**

Similarly, more lexical items are listed in LFq (short poem) than in the other two text types. Short poems, as mentioned above, do not prefer mundane words and prefer using poetry words or nature metaphors to render feelings and emotions into short poems. This strict diction thus excludes many lexical items from the text type of short poem. Among seven particles, three particles (*te*, *ni* and *do*) are classified as conjunctions and two particles (*tote* and *nado*) are classified as citational markers. 13 out of 27 verbs are classified as honorifics, either forms of respect, forms of humbleness or forms of politeness: *tamau* 'to indicate respect', *obosu* 'to think', *haberi* 'to show politeness', *kikoyu* 'to say', *tatematsuru* 'to give', *notamau* 'to say', *mairu* 'to go', *owasu* 'to exist', *owashimasu* 'to exist', *moosu* 'to say', *saburau* 'to serve', *tamau* 'to indicate humbleness', and *notamawasu* 'to say'. As for auxiliary verbs, the auxiliary verbs, *ri* and *tari*, indicate resultative, continuative or perfective while three auxiliary verbs, *nari*, *meri*, and *maji*, are classified as modals. The auxiliary verbs *su* and *sasu* indicate causative and honorific; the auxiliary *raru* indicates passive, honorific, spontaneous, and potential. Most of the nouns listed in LFq (short poem) are either social title, such as *otodo*, *chuujo*, and *taishoo*, abstract nouns, such as *koto* 'thing', *hodo* 'degree', and *tokoro* 'place', or words indicating people, such as *onna* 'woman', *hitobito* 'people', and

otoko ‘man’. Also, the adjective *okashi*, a representative word for Heian literature, is listed as one of the ten adjectives for LFq (short poem).

3.3.3. A Summary of Findings (1)

As seen above, certain linguistic tendencies can be found in each type of text. In the text types of narration and conversation, approximately 1.0% of the lexical items are considered as either highest frequency (i.e., preferred) words or lowest frequency (i.e., dispreferred) words. In contrast, in the text type of short poem, 2.8% of the lexical items are considered as highest frequency (i.e., preferred) words or lowest frequency (i.e., dispreferred) words. That is, there are more specialized lexical items in short poems compared to narration and conversation. Some specialized lexical items that are categorized as HFq (short poem) are lexical items that have been identified as *utakotoba*, which are preferably used in short poems.

Another finding was that complementary distribution of certain lexical items across different text types. Lexical items, especially particles, auxiliary verbs, verbs of honorifics and some affixes and nouns, are clearly distributed in a complementary manner between narrations and conversations. Lexical items that appear as HFq (narration) appear as LFq (conversation). Likewise, lexical items that appear as LFq (narration) appear as HFq (conversation). This finding statistically confirms Ijima (2011), which claims that not all linguistic features equally appear in every text type due to their functional differences. First, particles categorized as conjunction, such as *te*, *ni*, and *ba*, are listed in the HFq (narration) and in the LFq (conversation) categories whereas particles utilized for exclamation or question, such as *kana*, *ka* and *ya*, are listed in the LFq (narration) and the HFq (conversation) categories. Also, the particles classified as citational markers *to* and *tote* are distributed in a complementary fashion between narration and conversation, which are listed as HFq (narration) as well as LFq (conversation). These distributions are not surprising as

narration requires a series of clauses combined together to display temporally and causally connected events and facts, and to quote conversations of characters. On the other hand, speakers of conversations need to express their emotions through exclamation and interact with an interlocutor through a question and answer sequence. Secondly, though *keri* also indicates exclamatory recognition, the auxiliary verbs listed in HFq (narration) as well as in LFq (conversation), such as *tari*, *ri*, and *keri*, are categorized as tense and aspect; the auxiliary verbs listed in LFq (narration) as well as HFq (conversation), such as *mu*, *beshi* and *meri*, are categorized as modals. The auxiliary verbs *tari* and *ri* indicate resultive, continuative or perfective, while the auxiliary verb *keri* indicates hearsay past or exclamatory recognition. On the other hand, the auxiliary verbs such as *mu beshi* and *meri* refer to speculation and supposition. This distribution is also due to different functions between narration and conversation. Sentences of narration describe events and characters' actions, which indicate when the event/action takes place and show how the action or the event, such as its completion and progress, is viewed. Sentences of conversation, on the other hand, describe characters' thoughts, emotions, or feelings, which are to be expressed with the speaker's attitude toward the predication. Another complementary distribution is seen in verbs for honorifics. Forms of humbleness, such as *tamau* 'to indicate humbleness' and *makaru* 'to exit', and forms of politeness, such as *haberi* 'to indicate politeness', are listed in HFq (conversation) as well as in LFq (narration), while forms of respect, such as *obosu* (to think) and *notamau* (to say), are listed in HFq (narration) as well as in LFq (conversation). That is, forms of respect tend to appear in narration while forms of humbleness and forms of politeness tend to appear in conversation. This distribution of honorifics statistically confirms that the form of politeness *haberi*, which ranks first in HFq (conversation), and the form of humbleness *tamau*, which ranks fourth in HF (conversation), are especially utilized in conversation (Nishida, 1987; Oda, 2015). *Haberi* is used to indicate respect towards the addressee while *tamau* is used to humble the speaker in relation to the addressee. That is,

these lexical elements have communicative functions. There are other lexical items that show a complementary distribution between narration and conversation: lexical items that describe objects from outside and interjections. The affixes *ge* 'to appear to' and *garu* 'to show signs of' as well as the nouns *kehai* 'appearance,' *sama* 'appearance,' and *keshiki* 'appearance,' which describe objects from outside, are listed in HFq (narration) as well as LFq (conversation). In contrast, the interjections *ana* and *ide* are listed in LFq (narration) as well as in HFq (conversation). These sharp contrasts between narration and conversation also stem from the distinctive functions between narration and conversation. Sentences of narration provide a description of objects or characters such as appearance, which are described from outside, whereas sentences of conversation deliver the characters' exclamation of their feelings. As discussed so far, lexical items of highest frequency in each text type demonstrate linguistic characteristics that are suitable to the function of each text type. The text type of narration prefers lexical items of producing continuous sentences (e.g., particles for conjunctions such as *te*, *ni*, and *ba*) and lexical items of description (e.g., verbs that describe specific actions such as *iu*: to say, *kaku*: to write, and *naku*: to cry); however, these lexical items are not preferred in sentences of conversation. The text type of conversation prefers lexical items that have the communicative function (e.g., particles for question, such as *ka*, and for prohibitive, such as *na*) and the expressive function (e.g., interjections, such as *ide*, and particles for exclamatory, such as *kana*) to render communication between characters as well as characters' emotions and feelings.

The text type of short poem has its own linguistic preference to render poems. For example, lexical items categorized as poetry words are preferably employed in short poems, such as *hana* 'flower,' *tsuyu* 'dew' and *uguisu* 'Japanese bush warbler.' Also, this text type prefers lexical items to indicate nature, which are used to avoid direct expressions or mundane words. For example, the words *tsuyu* 'the dew' is used to describe the action of crying instead of the word *naku* 'to cry.' In addition, some lexical items classified as HFq

(short poem) are employed for techniques of short poems. For example, *matsu* is used for *kakekotoba* (pivot word), which is a word that has multiple meanings in the context; *matsu* can mean ‘to wait' and ‘pine tree.' Another example is the word *hisakata*, which is used for *makurakotoba* (epithet). *Makurakotoba* is a fixed expression that is placed prior to specific words to enhance the rhythm of a poem, yet the meaning is often lost. The *makurakotoba*, *hisakata-no*, usually precedes these words, such as *hikari* (light/shining), *tsuki* (moon), and *kumo* (cloud). On the other hand, some lexical items of HFq (short poem) are categorized with HFq (conversation), such as the particles for exclamatory *kana* and *ya*. These particles are preferably used in short poems as well as conversation to express the composer’s/ the speaker’s feelings and emotions.

3.3.4. Log Likelihood Ratio on the Text of *The Tale of Genji*

In 3.3.1 above, I reported my findings for the 16 literary classical texts. Across the three text types (narration, conversation, short poem), the total number of the word tokens of the CHJ of the Heian Period Series is 843,703 while the total number of the word token of *The Tale of Genji* is 443,352. That is, the data of *The Tale of Genji* consists of almost 50% of the entire CHJ of the Heian Period Series (see Table 3.3). This means that the LLR results acquired from the 16 literary works may have been affected by linguistic tendencies of the tale. This section, therefore, performs LLR analysis on the texts of *The Tale of Genji* as well as on the other 15 literary works separately, in order to confirm that the LLR results of the 16 literary works represent overall linguistic characteristics of Heian literary works including Genji. Figure 3.4. provided below shows the distribution of the LLR values of narration for *The Tale of Genji*. As before, the vertical axis indicates the numbers of lexical items while the horizontal axis indicates the values of LLR.

Figure 3.4. Distribution of the LLR Values for Narration of *The Tale of Genji*

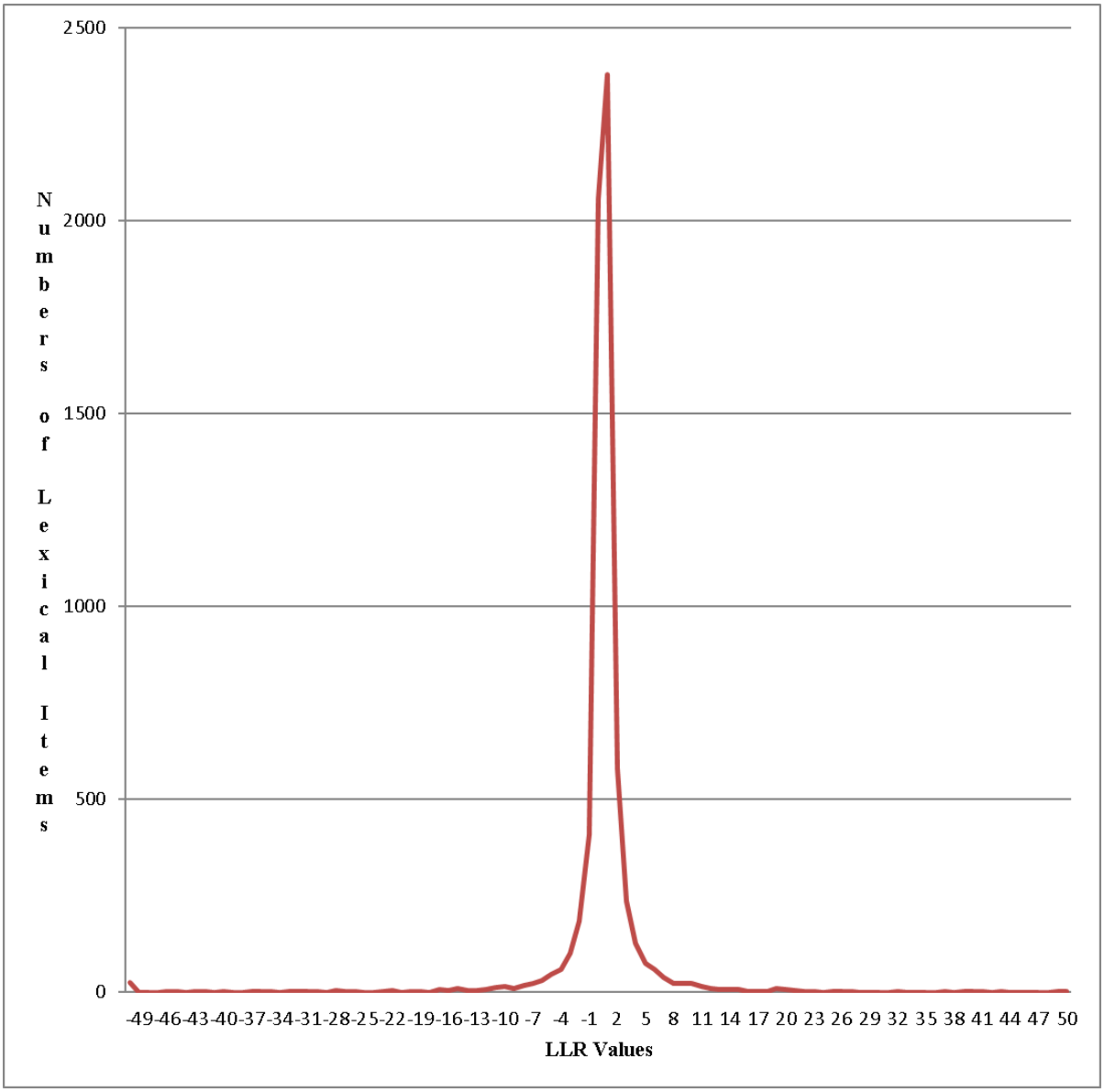


Figure 3.5. provided below shows the distribution of the LLR values of conversation for *The Tale of Genji*.

Figure 3.5. Distribution of the LLR Values for Conversation of *The Tale of Genji*

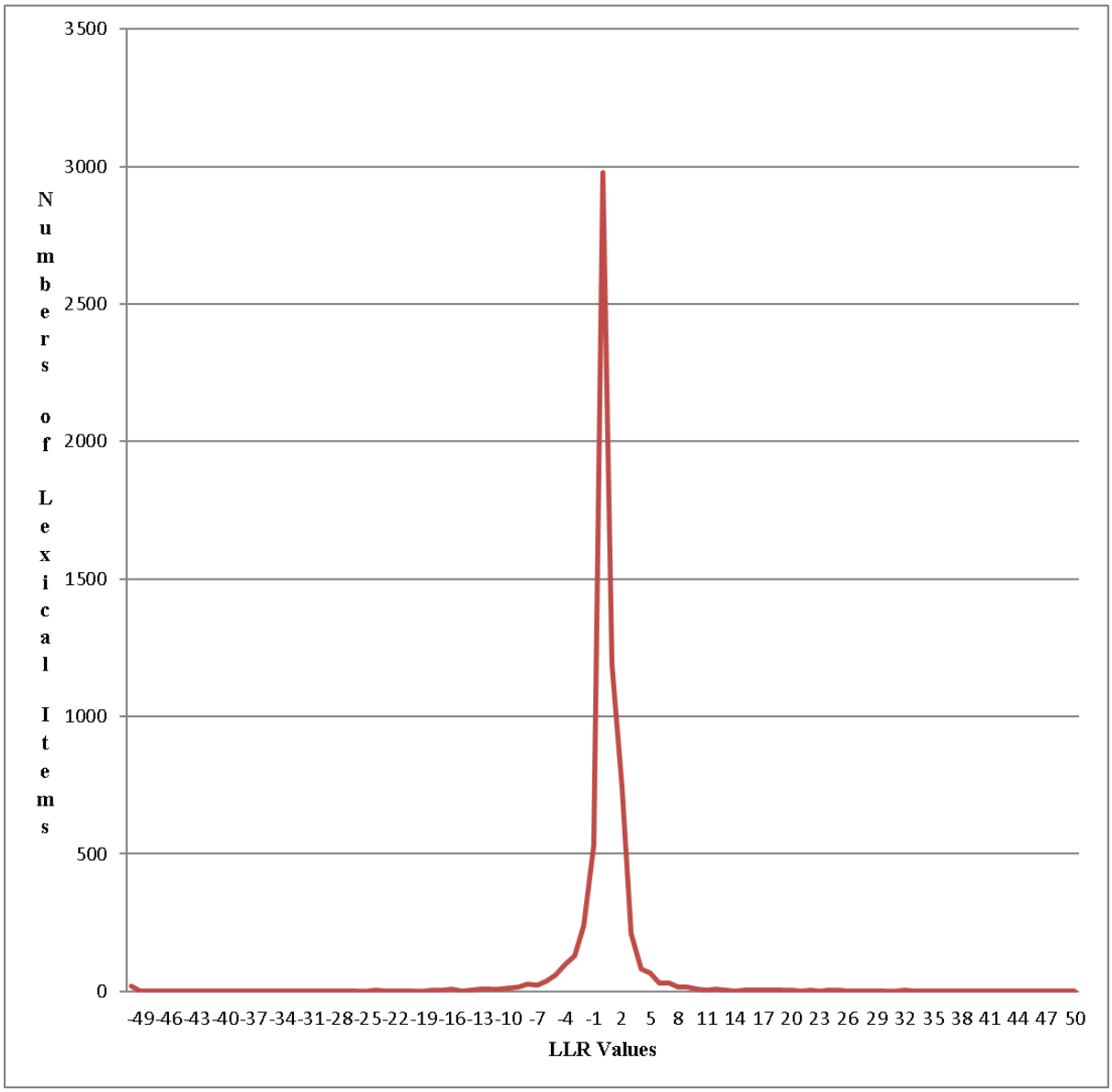
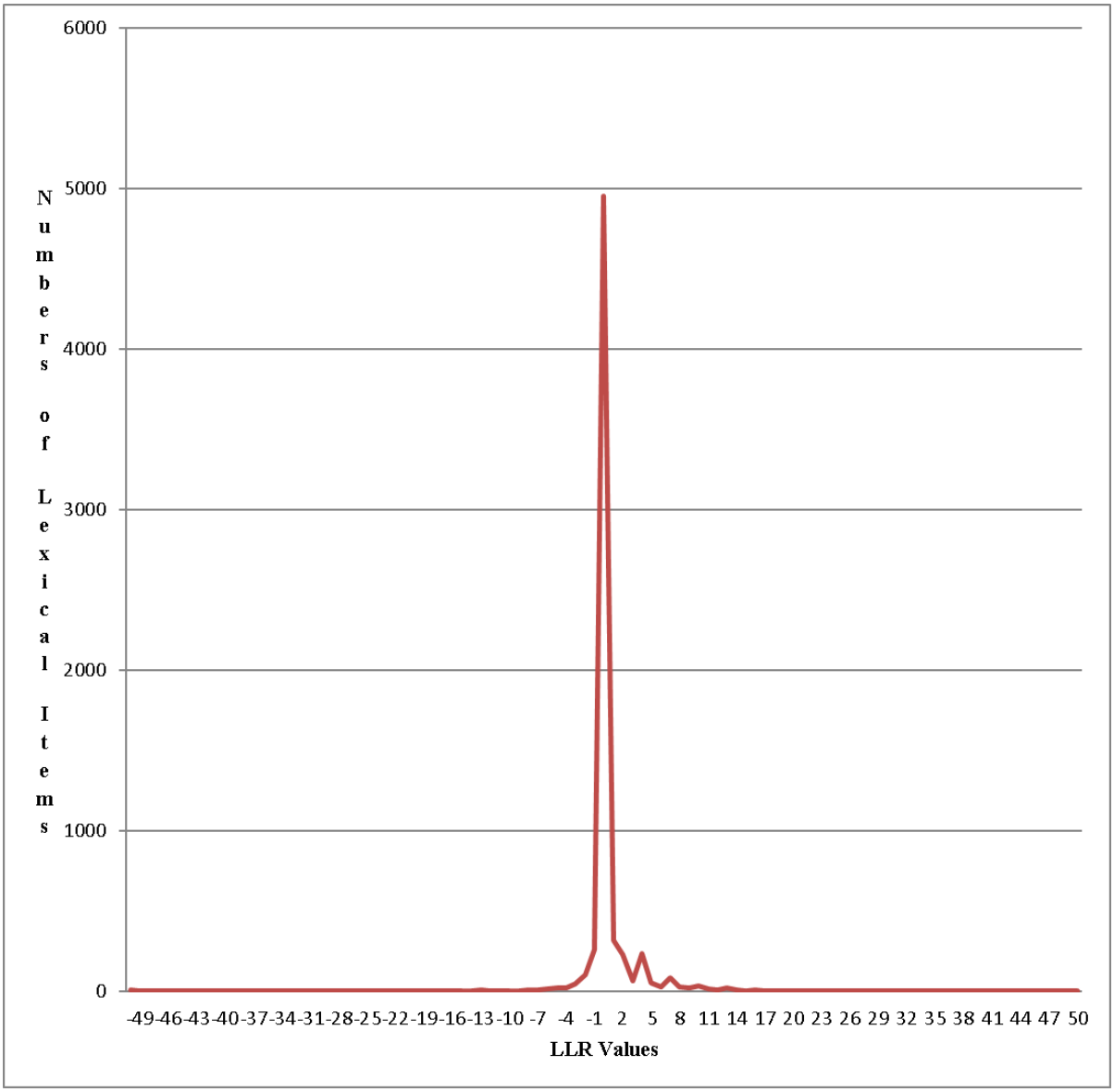


Figure 3.6. provided below shows the distribution of the LLR values of short poems for *The Tale of Genji*.

Figure 3.6. Distribution of the LLR Values for Short Poem of *The Tale of Genji*



As seen in each figure above, a large number of lexical items have LLR values that are located very close to 0.0 across the three text types, which is also seen in the distribution of LLR values in the 16 literary works.

Figure 3.7 provided below shows the distribution of LLR values of narration for the 15 literary works.

Figure 3.7. Distribution of the LLR Values for Narration of 15 literary works

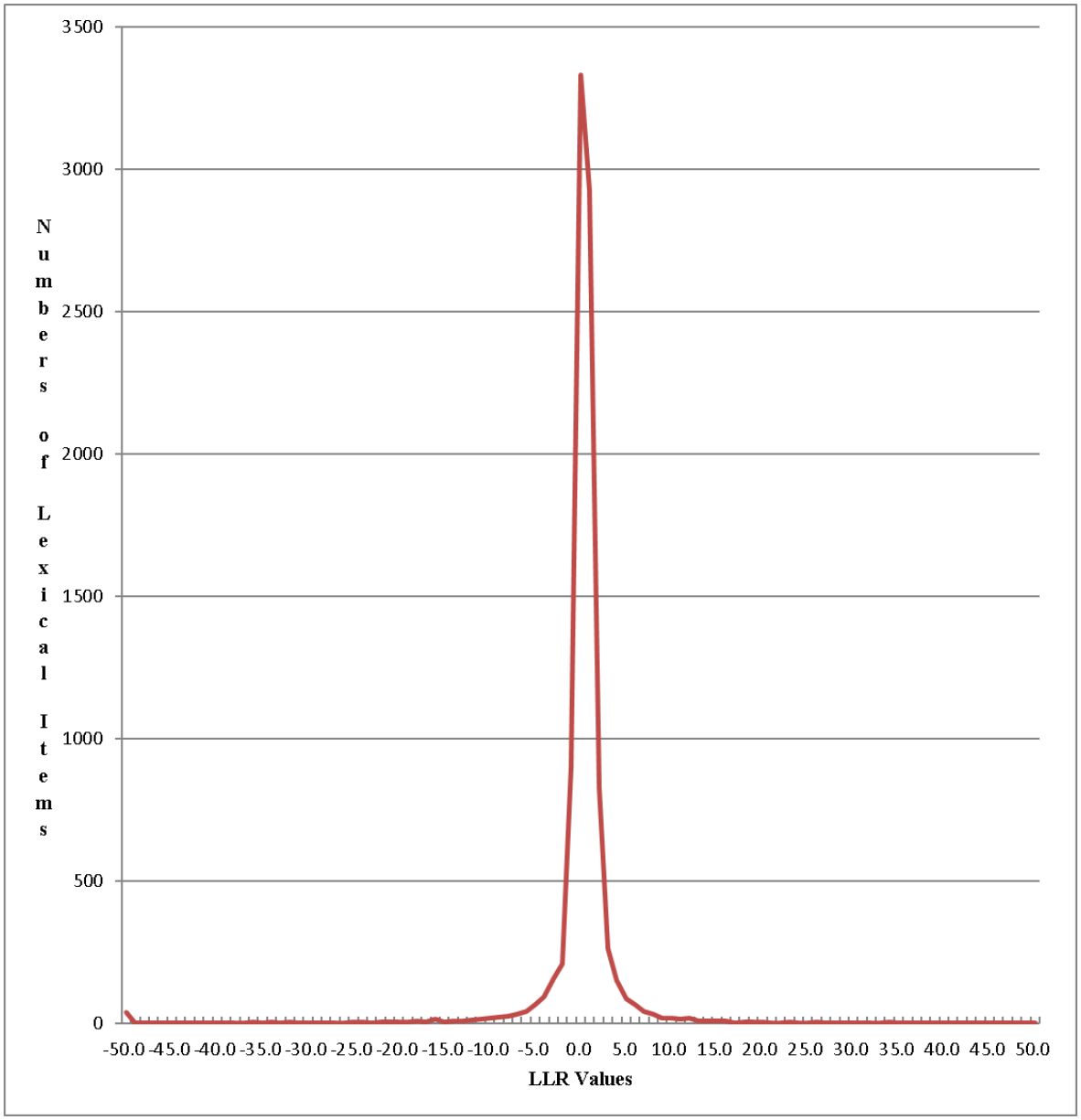


Figure 3.8. provided below shows the distribution of the LLR values of conversation for the 15 literary works.

Figure 3.8. Distribution of the LLR Values for Conversation of 15 literary works

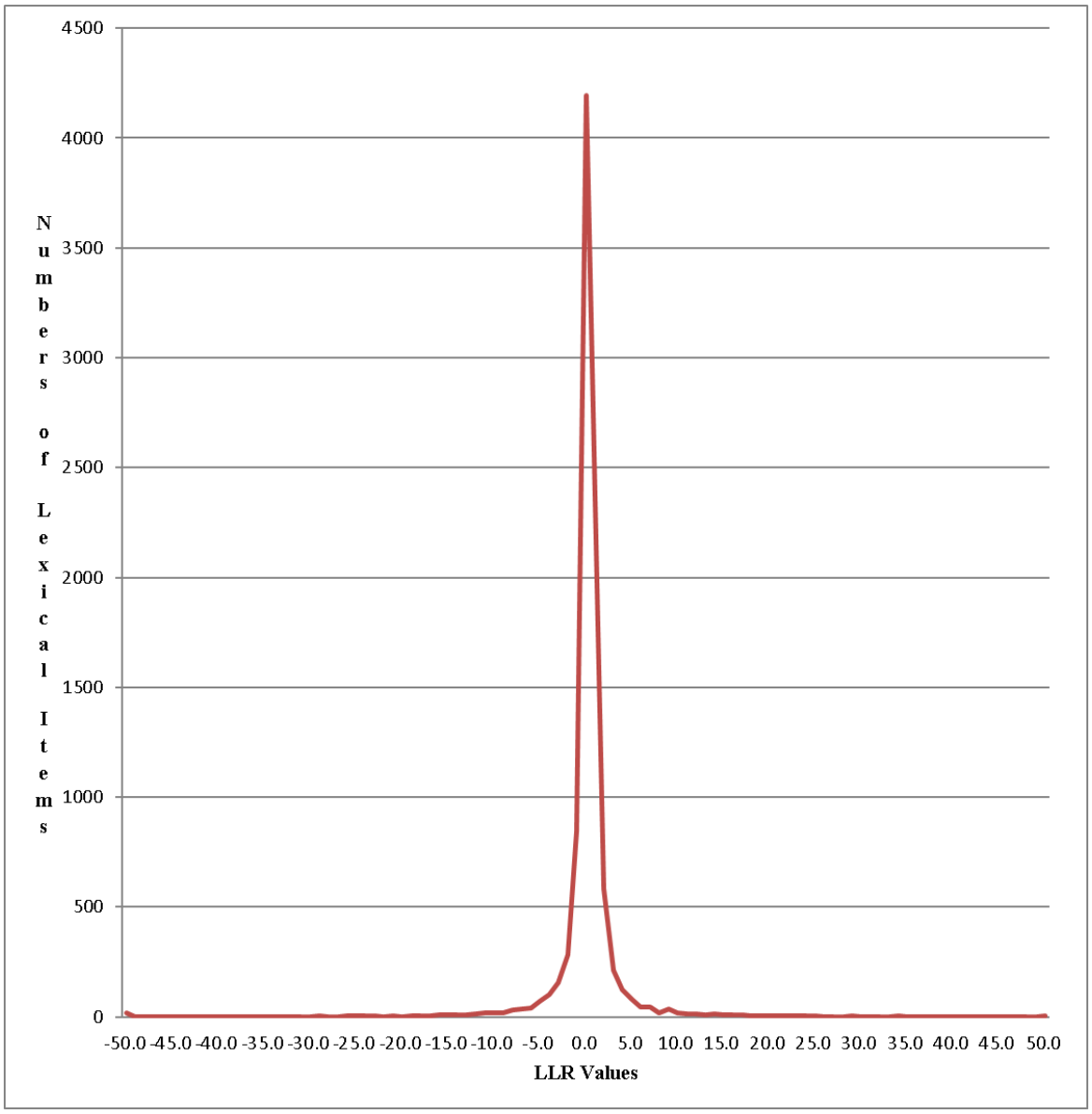
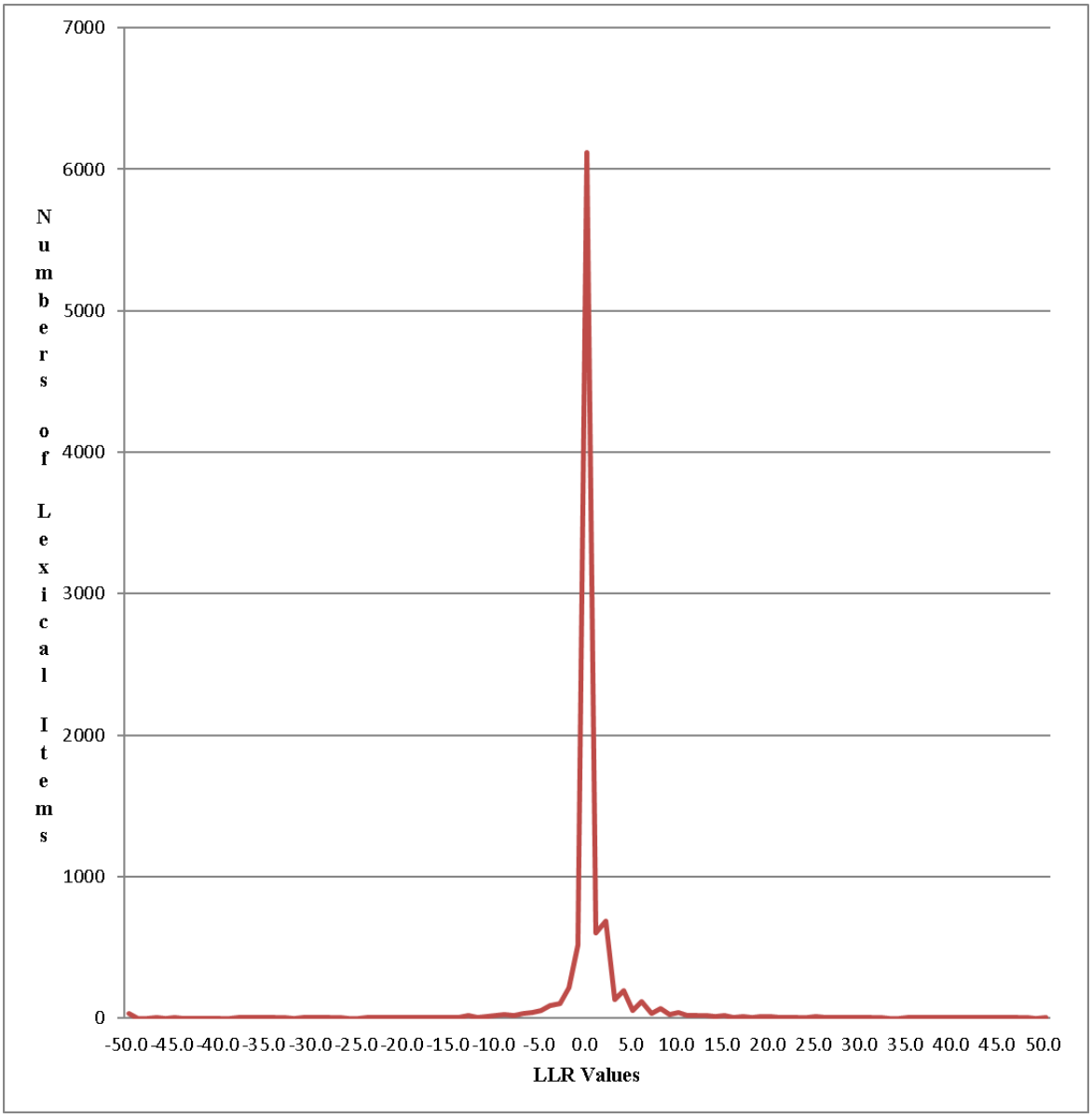


Figure 3.9. provided below shows the distribution of the LLR values of short poem for the 15 literary works.

Figure 3.9. Distribution of the LLR Values for Short Poem of 15 literary works



As seen in the last three figures, a large number of lexical items have LLR values located close to 0.0 across the three text types. This distribution pattern of LLR values is similar to the 16 literary works as well as in *The Tale of Genji* and suggests that most lexical items are almost equally employed in all the text types. As seen in the distribution of the 16 literary works, there are some lexical items both in *The Tale of Genji* and the 15 literary works that are spread out across different LLR values. In other words, some lexical items tend to appear more frequently or less frequently in one text type over the others.

3.3.5. Prominent Lexical Items in Different Text Types

In this section, I examine in more detail lexical items that are located widely spread out across different LLR values for the purpose of investigating what kind of lexical items show a prominent appearance in each text type for *The Tale of Genji* and the other 15 literary works. I consider the distribution of the values of LLR in each text type and utilize one standard deviation in the same manner as the 16 literary works. First, I investigated lexical items of narration both in *The Tale of Genji* as well as the 15 literary works. For narration of *The Tale of Genji*, the mean value is -0.44 and the standard deviation is 33.08; 99.0% of the lexical items of narration are included within the standard deviation. That is, 1% of the lexical items of narration are considered as either HFq (narration) or LFq (narration). Among these/ those lexical items, 32 lexical items are classified as HFq (narration). Provided below is the list of highest frequency words for narration across different parts-of-speech in *The Tale of Genji*.

Table 3.10. List of Preferred Words for Narration in *The Tale of Genji*

parts of speech	frequency
adjective	2
particle	7
auxiliary verb	2
prefix	2
suffix	3
verb	11
noun	5
Total	**32**

Lexical items classified as HFq (narration) do not have the communicative function or the expressive function, as seen in HFq (narration) for the 16 literary works. This linguistic tendency, as discussed before, is due to the function of the text type of narration; sentences of

narration function to unfold a story and describe characters and events. Out of seven particles, four particles (*te*, *ni*, *tsutsu*, and *ba*) are classified as conjunctions while three particles (*to*, *tote*, and *nado*) are classified as citational markers. The two auxiliary verbs listed are *ri* and *tari*, which indicate resultative, continuative, or perfective. Among 11 verbs, four verbs are classified as honorifics: *tamau* 'to indicate respect,' *obosu* 'to think,' *notamau* 'to say,' and *owasu* 'to exist,' all of which are forms of respect. The other verbs indicate specific actions of characters, such as *naku* 'to cry,' *iru* 'to sit,' and *kaku* 'to write.' There are two lexical items from different parts of speech that describe an object from outside: the suffix *ge* 'to appear to' and the noun *kehai* 'appearance.' The adjective *okashi*, which is a representative word for the Heian literature, is listed as one of the two adjectives for HFq (narration).

There are 36 lexical items classified as LFq (narration). Table 3.11 provided below is the list of lowest frequency words for narration across different parts-of-speech in *The Tale of Genji*.

Table 3.11. List of Dispreferred Words for Narration in *The Tale of Genji*

parts of speech	frequency
interjection	2
particle	9
auxiliary verb	9
pronoun	3
verb	6
adverb	2
noun	5
Total	**36**

Lexical items listed as LFq (narration) mostly have either the communicative function or the expressive function as seen in LFq (narration) of the 16 literary works. There are two interjections, *ana* and *ide*. Out of nine particles, one particle (*tomo*) is classified as conjunction whereas the other particles have either the expressive or the communicative

function. For example, the particle *kana* indicates exclamatory while the particle *ka* indicates question. As for auxiliary verbs, six auxiliary verbs are classified as modals: *mu*, *beshi*, *nari*, *ji*, *meri*, and *ramu*. The auxiliary verb *ki* indicates past whereas the auxiliary verbs *tsu* and *nu* indicate perfective. Four out of six verbs are classified as honorifics: *haberi* 'to show politeness,' *tamau* 'to indicate humbleness,' *makaru* 'to go,' and *uketamawaru* 'to receive.' All of them are either forms of humbleness or forms of politeness.

For narration of the 15 literary works, the mean value is -0.211 and the standard deviation is 15.978; 97.8% of the lexical items of narration are included within one standard deviation. That is, 2.2% of the lexical items of narration are considered as either HFq (narration) or LFq (narration). 79 lexical items are classified as HFq (narration). Table 3.12 shows a/ the distribution of the highest frequency words for narration across different parts-of-speech in the 15 literary works.

Table 3.12. List of Preferred Words for Narration in 15 Literary Works

parts of speech	frequency
adjective	3
particle	6
auxiliary verb	2
prefix	2
suffix	4
verb	30
adverb	29
noun	29
attributive word	1
Total	**106**

More lexical items (106 lexical items) are listed in HFq (narration) for the 15 literary works than Genji (32 items). Though more lexical items are listed, those items classified in narration do not have communicative or expressive functions. Instead, they contribute to producing a continuous sentence or describing characters and events. Out of 6 particles, three particles (*te*,

ba, and *ni*) are used as conjunction while the other three particles (*nado*, *to* and *tote*) function as citational markers. Also, two suffixes (*ge* and *garu*) are utilized to describe an object from outside. There are 36 verbs, which all describe specific actions of characters, such as *iu* 'to say,' *kaku* 'to write' and *warau* 'to laugh'; however, there are no honorific verbs listed.

136 lexical items are classified as LFq (narration). Provided below is the list of lowest frequency words for narration across different parts-of-speech in the 15 literary works.

Table 3.13. List of Dispreferred Words for Narration in 15 Literary Works

parts of speech	frequency
interjection	3
adjective	3
particle	20
auxiliary verb	14
prefix	4
suffix	5
pronoun	11
verb	23
adverb	8
noun	45
Total	**136**

As opposed to the lexical items of HFq (narration), there are lexical items that have the communicative function or the expressive function. There are three interjections listed (*ana*, *iza*, and *ide*) and particles utilized for exclamatory, such as *ya* and *kana*. Also, nine out of 14 auxiliary verbs are classified as modals: *mu*, *ramu*, *ji*, *maji*, *nari* (hearsay), *beshi*, *meri*, *rashi*, and *beranari*. As for verbs, 12 out of 23 verbs are classified as verbs of honorifics including forms of respect (e.g., *tamau* and *owashimasu*) forms of humbleness (e.g., *tamau* and *moosu*), and forms of politeness (e.g., *haberi*).

For conversation of *The Tale of Genji*, the mean value is 0.44 and the standard deviation is 34.76; 99.0% of the lexical items are included within the standard deviation. That is, 1% of the lexical items of conversation are considered to be either HFq (conversation) or

LFq (conversation). 35 lexical items are classified as HFq (conversation). Table 3.14 shows the distribution of the highest frequency words for conversation in *The Tale of Genji*.

Table 3.14. List of Preferred Words for Conversation in *The Tale of Genji*

parts of speech	frequency
interjection	2
particle	9
auxiliary verb	10
pronoun	2
verb	6
adverb	2
noun	4
Total	**35**

Lexical items classified as HFq (conversation), as seen before, have the communicative function or the expressive function, which are suitable to the function of the text type of conversation. Out of nine particles, one particle (*tomo*) is classified as conjunction while other particles have either the expressive or the communicative functions. For instance, the particles *ya* and *kana* indicate exclamatory and the particle *na* indicates prohibitive in conjunction with the particle *so*. Six out of ten auxiliary verbs are classified as modals: *mu*, *beshi*, *nari*, *meri*, *ji*, and *maji*. The auxiliary verb *ki* indicates past and the auxiliary verb *sasu* indicates causative or honorific while the auxiliary verbs *tu* and *nu* indicate perfective. As for verbs, four verbs are classified as honorifics: *haberi* ‘to show politeness,' *tamau* ‘to indicate humbleness,' *uketamawaru* ‘to receive,' and *makaru* ‘to go.' All of them are either forms of humbleness or forms of politeness.

There are 28 lexical items classified as LFq (conversation). Provided below is the list of lowest frequency words for conversation across different parts-of-speech in *The Tale of Genji*.

Table 3.15. List of Dispreferred Words for Conversation in *The Tale of Genji*

parts of speech	frequency
adjective	2
particle	7
auxiliary verb	2
prefix	2
suffix	3
verb	9
noun	3
Total	**28**

As opposed to the lexical items of HFq (conversation), lexical items listed as LFq (conversation) does not have the communicative function or the expressive function. As for particles, four particles (*te*, *tsutsu*, *ni*, and *ba*) are classified as conjunctions while three particles (*to*, *tote*, and *nado*) are utilized as citational markers. The two auxiliary verbs listed are *ri* and *tari*, which indicate resultative, continuative, or perfective. Three out of nine verbs are classified as honorifics: *tamau* ‘to indicate respect,' *obosu* ‘to think,' and *notamau* ‘to say.' All of the honorific verbs are forms of respect. The other verbs describe specific action, such as *naku* ‘to cry,' *iru* ‘to sit’ and *izu* ‘to go out.' There are two lexical items from different parts of speech that describe an object from outside: the suffix *ge* ‘to appear to’ and the noun *kehai* ‘appearance.' Also, listed is the adjective *okashi*, which is a representative word for the Heian literature.

For conversation of the 15 literary works, the mean value is -0.354 and the standard deviation is 15.207; 97.6% of the lexical items of narration are included within the standard deviation. That is, 2.4% of the lexical items of narration are considered as either HFq (conversation) or LFq (conversation). Among these lexical items, 128 lexical items are classified as HFq (conversation). Provided below is the list of highest frequency words for conversation in the 15 literary works.

Table 3.16. List of Preferred Words for Conversation in 15 Literary Works

parts of speech	frequency
interjection	2
adjective	5
particle	11
auxiliary verb	11
prefix	4
suffix	6
pronoun	10
verb	22
adverb	10
noun	47
Total	**128**

More lexical items are classified as HFq (conversation) than the 16 literary works or *The Tale of Genji*, yet the same linguistic tendencies are identified. Lexical items that have the communicative function or the expressive function are preferred. There are particles that have the communicative function or the expressive function. For example, the particle *na* is utilized for prohibitive in conjunction with the particle *so*, and the particle *kana* is utilized for exclamatory. Also, there are two interjections listed: *ana* and *ide*. Out of 11 auxiliary verbs, four auxiliary verbs are classified as modals: *mu*, *meri*, *beshi* and *nari* (hearsay). As for verbs, 13 out of 22 verbs are classified as honorific verbs including forms of respect (e.g., *tamau* and *owasu*), forms of humbleness (*tatematsuru* and *moosu*) and forms of politeness (*haberi* and *saburau*). 101 lexical items are classified as LFq (conversation). Table 3.17 shows the distribution of the lowest frequency words for conversation across different parts-of-speech in the 15 literary works.

Table 3.17. List of Dispreferred Words for Conversation in 15 Literary Works

parts of speech	frequency
adjective	3
particle	11
auxiliary verb	2
prefix	2
suffix	4
pronoun	2
verb	45
noun	31
attributive word	1
Total	**101**

In contrast to the lexical items of HFq (conversation), lexical items of LFq (conversation) do not have either the communicative function or the expressive function. Out of 11 particles, four particles (*te*, *ba*, *ni*, *and tsutsu*) are classified as conjunction and three particles (*nado*, *to*, and *tote*) function as citational markers. As for auxiliary verbs, two auxiliary verbs listed are *tari*, which indicates resultative, continuative or perfective, and *keri*, which indicates hearsay past or exclamatory recognition, while there are no auxiliary verbs classified as modals. Though there are 45 verbs listed no verbs are classified as honorific verbs. Also, there are two suffixes to describe an object from outside: the suffixes *ge* 'to appear to' and *garu* 'to show signs of.'

For short poems of *The Tale of Genji*, the mean value is 0.36 and the standard deviation is 7.57; 95.9% of the lexical items are included within the standard deviation. That is, 4.1% of the lexical items of short poems are considered to be either HFq (short poem) or LFq (short poem). 199 lexical items are classified as HFq (short poem). Provided below is the list of highest frequency words for short poem across parts-of-speech in *The Tale of Genji*.

Table 3.18. List of Preferred Words for Short Poem in *The Tale of Genji*

parts of speech	frequency
adjective	5
particle	14
auxiliary verb	6
suffix	1
pronoun	3
verb	58
noun	112
Total	**199**

The text type of short poem, as discussed before, has a preference for lexical items known as poetry words as well as those of nature. Therefore, a large number of lexical items of HFq (short poem) are listed as poetry words. Also, there are lexical items that have the expressive function to convey the composer's feelings. As for particles, some particles have the expressive or the communicative functions, such as *kana* for exclamation and *ka* for question, while other particles are classified as conjunction, such as *domo* and *tomo*. Four out of six auxiliary verbs are modals: *ramu*, *rashi*, *mashi*, and *ji*. The auxiliary verb *ki* indicates past while the auxiliary verb *zu* indicates negative. Out of 58 verbs, no verbs are classified as honorifics, either forms of respect, forms of humbleness, or forms of politeness. There are nouns listed in HFq (short poem) more excessively than the other categories. Most of the nouns may be categorized into nature, clothes, plants, and animals. Into the category of nature are categorized nouns such as *tsuyu* 'dew,' *shimo* 'frost,' and *kiri* 'mist' while into the category of clothes are categorized nouns such as *sode* 'sleeves,' *natsu goromo* 'summer clothes,' and *tamoto* 'sleeves.' For plants are included nouns such as *matsu* 'pine trees,' *nadeshiko* 'dianthus flower,' and *kiku* 'chrysanthemum'; for animals are included nouns such as *uguisu* 'Japanese bush warbler,' *tazu* 'crane,' and *hototogisu* 'lesser cuckoo.'

There are 77 lexical items classified as LFq (short poem). Provided below is the list of LFq (short poem) across different parts-of-speech.

Table 3.19. List of Dispreferred Words for Short Poem in *The Tale of Genji*

parts of speech	frequency
adjective	7
particle	8
auxiliary verb	10
prefix	3
suffix	4
pronoun	1
verb	21
adverb	7
noun	16
Total	**77**

In contrast to the lexical items of HFq (short poem), lexical items listed as LFq (short poem) are not suitable for short poems. Two particles (*nado* and *do*) are classified as citational markers while four particles (*te*, *ni*, *ba*, and *do*) are classified as conjunctions. Out of ten auxiliary verbs, four auxiliary verbs are classified as modals: *meri*, *maji*, *nari*, and *beshi*. Two auxiliary verbs (*ri* and *tari*) indicate resultative, continuative or perfective; two auxiliary verbs (*su* and *sasu*) indicate causative, passive and honorific; the auxiliary verb *raru* indicates passive, honorific, spontaneous, and potential. 12 out of 21verbs are classified as honorifics, either forms of honorifics, forms of humbleness, or forms of politeness: *tamau* 'to indicate respect', *obosu* 'to think', *kikoyu* 'to say', *haberi* 'to show politeness', *tatematsuru* 'to give', *notamau* 'to say', *mairu* 'to go', *owasu* 'to exist', *owashimasu* 'to exist', *tamau* 'to indicate humbleness', *saburau* 'to serve', and *moosu* 'to say'. Most nouns are categorized into social a title such as *in*, *taishoo*, and *otodo*, people such as *hitobito* 'people' and *onna* 'woman,' or abstract noun such as *koto* 'thing,' *sama* 'condition,' and *hodo* 'degree.' Also, *okashi*, a representative word in the Heian literature is listed as LFq (short poem).

For short poem of the 15 literary works, the mean value is -0.366 and the standard

deviation is 15.350; 96.3% of the lexical items of narration are included within the standard deviation. That is, 3.7% of the lexical items of narration are considered as either HFq (narration) or LFq (narration). 221 lexical items are classified as HFq (narration). Table 3.20 shows the distribution of the highest frequency words for short poem across different parts-of-speech in the 15 literary works.

Table 3.20. List of Preferred Words for Short Poem in 15 Literary Works

parts of speech	frequency
adjective	10
particle	21
auxiliary verb	8
suffix	2
pronoun	4
verb	53
noun	123
Total	**221**

Lexical items of HFq (short poem) for the 15 literary works show the same linguistic tendencies as those for the 16 literary works as well as *The Tale of Genji*. Most of them are known as poetry words or related to nature. There are lexical items of HFq (short poem) listed as poetry words and lexical items that have the expressive function to convey the composer's feelings. As for particles, some particles have the expressive or the communicative functions, such as *kana* and *ya* for exclamation and *ka* for question, while other particles are classified as conjunction, such as *domo*, *o*, and *tsutsu*. Seven out of eight auxiliary verbs are modals: *ramu*, *zu*, *mashi*, *ji*, *rashi*, *beranari*, and *mu*. Out of 53 verbs, no verbs are classified as honorifics, either forms of respect, forms of humbleness, or forms of politeness. There are nouns listed in HFq (short poem) more excessively than the other categories. Most of the nouns may be categorized into nature, clothes, plants, and animals. Into the category of nature are categorized nouns such as *tsuyu* 'dew,' *shimo* 'frost,' and

shigure ‘shower’ while into the category of clothes are categorized nouns such as *sode* ‘sleeves,' *koromo* ‘clothes,' and *tamoto* ‘sleeves.’ For plants are included nouns such as *hana* ‘flower,' *momiji* ‘Japanese maple,' and *sakurabana* ‘cherry blossom’; for animals are included nouns such as *uguisu* ‘Japanese bush warbler,' *shika* ‘deer,' and *hototogisu* ‘lesser cuckoo.’ 132 lexical items are classified as LFq (narration). Provided below is the list of lowest frequency words for short poem across parts-of-speech in the 15 literary works.

Table 3.21. List of Dispreferred Words for Short Poem in 15 Literary Works

parts of sppech	frequency
adjective	7
particle	6
auxiliary verb	7
conjunction	2
prefix	6
suffix	7
pronoun	5
verb	39
adverb	7
noun	46
Total	**132**

Lexical items of LFq (short poem) also show the same linguistic tendencies as for the 16 literary works as well as *The Tale of Genji*. There are two particles (*nado* and *do*) classified as citational markers while two particles (*te*, and *ni*) classified as conjunctions. Out of seven auxiliary verbs, two auxiliary verbs are classified as modals: *meri* and *maji*. One auxiliary verb (*tari*) indicates resultative, continuative or perfective; two auxiliary verbs (*su* and *sasu*) indicate causative, passive and honorific; the auxiliary verb *raru* indicates passive, honorific, spontaneous, and potential. The auxiliary verb *nari* indicates declarative. 16 out of 39verbs are classified as honorifics, either forms of honorifics, forms of humbleness, or forms of politeness, such as *tamau* ‘to indicate respect,' *obosu* ‘to think,' and *kikoyu* ‘to say.’ Most

nouns are categorized into social title such as *in*, *shooshoo*, and *otodo*, people such as *otoko* ‘man’ and *onna* ‘woman,' or abstract noun such as *koto* ‘thing,' *sama* ‘condition,' and *hodo* ‘degree.' Also, *okashi*, a representative word in the Heian literature is listed as LFq (short poem).

3.3.6. A Summary of Findings (2)

Section 3.3.5. discussed the linguistic tendencies of the three text types (narration, conversation, and short poem) that appear in *The Tale of Genji* as well as in the other 15 literary works. For *The Tale of Genji*, in the text types of narration and conversation, approximately 1.0% of the lexical items are considered as either highest frequency or lowest frequency words whereas in the text type of short poem, 4.1% of the lexical items are considered as highest or lowest frequency words. For the 15 literary works, on the other hand, in the text types of narration and conversation, approximately 2.3% of the lexical items are considered as either highest or lowest frequency words whereas in the text type of short poem, 3.7% of the lexical items are considered as highest or lowest frequency words. That is, more lexical items tend to appear in short poems over the other text types due to the characteristic of the diction of short poems. As for the text types of narration and conversation, there is a complementary relationship between HFq (narration) and LFq (conversation) as well as LFq (narration) and HFq (conversation) due to the function of each text types; sentences of narration dedicated to describing characters and producing continuous sentences to unfold a story while sentences of conversation function to demonstrate communication between characters and to express characters feeling. Particles categorized as conjunction, such as *te*, *ni*, and *ba*, are listed in HFq (narration) as well as in LFq (conversation) whereas particles utilized for exclamation or question, such as *kana*, *ka* and *ya*, are listed in LF (narration) as well as HFq (conversation). The particles classified as citational markers *to* and *tote* are also listed in HFq (narration) as well as in LFq

(conversation). As for auxiliary verbs, the auxiliary verbs listed in HFq (narration) as well as in LFq (conversation) are the auxiliary verbs *tari*, *ri*, and *keri*. The auxiliary verbs *tari* and *ri* indicate resultative, continuative or perfective, while the auxiliary verb *keri* indicates hearsay past or exclamatory recognition. On the other hand, auxiliary verbs classified as modals are listed in LFq (narration) as well as HFq (conversation), such as *mu*, *beshi*, and *meri*. There are other lexical items that show a complementary distribution: lexical items that describe objects from outside and interjections. The affixes *ge* 'to appear to' and the noun *kehai* 'appearance,' which describe objects from outside, are listed in HFq (narration) as well as LFq (conversation). In contrast, the interjections *ana* and *ide* are listed in LFq (narration) as well as in HFq (conversation). Honorific verbs, on the other hand, show a different linguistic tendency for the 15 literary works. For the 15 literary works, verbs of honorifics are classified as LFq (narration) as well as HFq (conversation). For the 16 literary works and *The Tale of Genji*, however, forms of respect are listed in HFq (narration) as well as LFq (conversation) whereas forms of humbleness and forms of politeness are listed in LFq (narration) as well as HFq (conversation). Therefore, honorific verbs, especially forms of respect, may have a prominent use in the texts of *The Tale of Genji*.

Except for verbs of honorifics, the linguistic tendencies identified in the texts of *The Tale of Genji* as well as the 15 literary works show overall similarity to those found in the texts of the 16 literary works. The text type of conversation prefers lexical items that have the communicative function or the expressive function while the text type of narration does not. For example, the auxiliary verbs *tari* and *ri*, which indicate resultative, continuative or perfective, are listed in HFq (narration) as well as LFq (conversation), while auxiliary verbs classified as modals are listed in LFq (narration) as well as HFq (conversation). Also, particles for conjunction, such as *to* and *tote*, are preferred in the text type of narration while particles for exclamatory, such as *kana* and *ya*, are preferred in the text type of conversation. On the other hand, the text type of short poem, which has a preference for the diction of short

poems, prefers lexical items of poetry words and lexical items indicating nature, such as *ume* 'Japanese apricot/ plum,' *shika* 'deer' and *tsuyu* 'dew.' In sum, these linguistic tendencies found in the current study suggest that there are certain linguistic tendencies in sentences of each text type. Also, the texts of *The Tale of Genji* were overall written in parallel with general linguistic tendencies identified in the texts of the other literary works from the Heian period.

3.4. Subjective Lexical Items in Conversation

3.4.1. Auxiliary Verbs of Modal Information

In HFq (conversation) are listed the auxiliary verbs, *ki*, *su*, *sasu*, *nu*, *mu*, *beshi*, *nari*, *meri*, *ji*, and *maji*, which in sharp contrast, do not appear in the text of narration. Provided below is the list of auxiliary verbs of classical Japanese. There are different ways to categorize the auxiliary verbs, such as categorization by meaning and categorization by conjugation type, yet I have employed the most common categorization, which is meaning based categorization (Kitahara, 1981).[25] Some auxiliary verbs are listed in multiple categories due to the nature of polysemy.[26]

[25] There are two different auxiliary verbs *nari*, which conjugate differently.

[26] The affirmative *nari* and the hearsay *nari* are not polysemy. These two *nari* conjugate differently, so they are different auxiliary verbs.

Table 3.22. List of Auxiliary Verbs of Classical Japanese

Declarative	Affirmative	nari, tari
	Negative	zu
Tense/ Aspect	Perfective	tsu, nu, tari, ri
	Past	ki, keri
Voice	Passive	ru, raru
	Causative	su, sasu, shimu
	Potential	ru, raru
	Spontaneous	ru, raru
Evidential/ Epistemic	Hearsay	nari
	Speculative	mu, ramu, meri, rashi, mashi, beshi, nari, kemu
	Negative speculative	ji, maji
Subjective	Desiderative	tashi, mahoshi
	Exclamative	keri
Comparative		gotoshi, gotoku-nari, yoo-nari
Treatment expression		ru, raru, su, sasu, shimu

Among the auxiliary verbs of HFq (conversation), the auxiliary verbs *su* and *sasu* indicate causative and honorific (labeled as 'treatment expression' in Table 3.22 above) while the auxiliary verb *nu* indicates perfective. On the other hand, the other auxiliary verbs of HF (conversation), *mu*, *beshi*, *nari*, *meri*, *ji*, and *maji*, are associated with subjectivity. The auxiliary verbs *mu*, *meri*, *beshi* indicate speculative and the auxiliary verbs *ji* and *maji* indicate negative speculation while the auxiliary verb *nari* indicates hearsay. The conjecture auxiliary verbs *mu*, *meri*, *beshi*, *ji*, *maji*, and *nari*, Kinsui (2011) claims, are categorized as epistemic modality (p.57). Epistemic modality indicates how much certainty or evidence the speaker has for the proposition expressed. In other words, epistemic modality shows the speaker's subjectivity. Provided below are example sentences in which these lexical items of modality classified as HFq (conversations) are utilized. Those auxiliary verbs that are associated with subjectivity are in bold.

(1) かぐや姫　の　皮衣　を　見　て　いはく
kaguyahime　no　kawagoromo　o　mi　te　iwaku
N (name)　P (case)　N (fur garment)　P (case)　V: to look at　P (conjective)　V: to say

「うるはしき　皮　な　**めり**　」
uruwashiki　kawa　na　**meri**
Adj: fine　N: fur　AUX (declarative)　**AUX: supposition based on visual evidence**

Looking at the fur garment, Princess Kaguya said "It looks like like beautiful fur." (*The Bamboo Cutter*)

(2) 僧都　は　よも　さやうに　は
soozu　wa　yomo　sayooni　wa
N:monk　P (bound)　Adv: rarely　Adj V: like that　P (bound)

据ゑ　たまは　**じ**　を
sue　tamawa　**ji**　o
V: to let live　V (respect)　**Aux: negative speculation**　P (case)

The monk cannot possibly have (women) live there like that. (*The Tale of Genji*)

(3) しづまり　ぬ　**なり**
shizumari　nu　**nari**
V: to become quiet　Aux (perfective)　**Aux: auditory speculation**

It seems like that everything has become quiet. (*The Tale of Genji*)

(4) いみじく　思し　嘆く　こと　ある　**べし**
imijiku　oboshi　nageku　koto　aru　**beshi**
Adj:extremely　V: to think　V: to lament　N: thing　V: to exist　**Aux: conjecture with confidence**

There must be something to think and lament over. (*The Bamboo Cutter*)

3.4.2. Interjections and Particles for Exclamation

The other targeted lexical items are interjections, such as *ana* and *ide*, and particles of exclamation, such as *kana* and *ya*. These lexical items are also associated with subjectivity. Scholars such as Lyons, Traugott, and Dasher, claim that subjectivity is not something just

encoded in linguistic expressions per se, but something added by the speaker in various ways, such as prosodically, when the sentence is uttered. However, associating subjectivity with various kinds of linguistic expressions and features, Iwasaki (1993, 2013) classifies these linguistic expressions and features associated with subjectivity into different categories: internal state expressions, deictic expressions, and movement and transaction expressions. Among different types of internal expressions proposed in Iwasaki (2013) are exclamatory sentences, which directly express the subjectivity of the speaker, such as perceptions, emotions, and feelings. In classical Japanese, the particles classified as HFq (conversation) *ka*, *ya*, *kana* and *yo*, which shows a sharp contrast between the text type of narration and the text type of conversation, indicate exclamation. That is, these particles directly express the subjectivity. Provided below is a sentence in which *kana* one of the particles classified as HFq (conversation) is utilized.

(5) なかなか 長き より も こよなう いまめかしき もの **かな**
nakanaka nagaki yori mo koyonoo imamekashiki mono **kana**
Adv: rather tha Adj: long P (case) P (bound) Adj: more Adj: modern N (thing) **P (final)**
(The nun's hair being short) is much more modern rather than being long. (*The Tale of Genji*)

Interjections, Iwasaki states, are lexical items that may also occur in exclamatory sentences to "clarify the exclamatory status of the sentence" (p. 253).[27] In other words, interjections are also used to express subjectivity. Provided below is a sentence where two interjections, *ide* and *ana*, are used in conjunction with *ya* one of the particles classified as HFq (conversation).

[27] Iwasaki (2013) calls this type of linguistic elements exclamatory vocal sign.

(6) **いで** **あな** 幼 **や**

ide **ana** osana **ya**

Intj **Intj** Adj: childish **P (bound)**

Oh my, how childish you are! (*The Tale of Genji*)

These lexical items discussed so far (auxiliary verbs categorized as epistemic modality, interjections, and particles for exclamation) are all associated with subjectivity and indicate the speaker's subjective involvement in the proposition. However, they are not associated with the communicative act. This type of lexical items are the ones proposed in the previous studies, such as Banfield (1982) and Kuroda (1973/ 2014), that make sentences read as if they were written in the first-person. Though linguistic elements that are identified to render this effect may vary across languages, Banfield and Kuroda claim that lexical items associated with subjectivity contribute to creating sentences in question. The current study, therefore, further investigates how these lexical items associated with subjectivity are utilized to create the unique effect in the texts of *The Tale of Genji*.

3.5. Quantitative Analysis 2: Concordance Plots

This section reports the results of a concordance search of the subjective lexical items that are described in the previous section (3.4). These lexical items, as the LLR analysis reveals, are all associated with subjectivity and not with communicative acts. The concordance search is performed on sentences of narration occurring throughout *The Tale of Genji* to examine where these subjectivity elements occur. Specifically focusing on selected passages known as *kaimami* 'peering', where a male character covertly peers upon a female character, the current concordance search provides concordance plots to visually show where the subjective lexical items occur.

3.5.1. *Kaimami* (peering)

Kaimami, the literal meaning of which is the action of peering through the gap, conventionally refers to the peering in which a man covertly sees a woman.[28] The word *kaimami* does not frequently appear in texts. However, the action of peering appears throughout anecdotes and tales of classical Japanese literature from different time periods. The action of peering appears in literature from the Nara period, such as *Records of Ancient Matters* (*Kojiki*) and *The Chronicles of Japan* (*Nihonshoki*). In the early Heian period, the word *kaimami* first appears in the 10th century text of *The Tale of Bamboo Cutter*, which is known as the oldest prose folktale (Murofushi, 1995; Yoshikai, 2008). *Kaimami* scenes, Yoshikai states, appear most often in the genre of tales, such as *The Tales of Ise* (*Ise Monogatari*), *The Tale of Ochikubo* (*Ochikubo Monogatari*), and *The Tale of the Hollow Tree* (*Utsubo Monogatari*). Across different time periods, in the conventional form of *kaimami*, the observer is a male character and the observed is a female character; a male character covertly sees a female character.[29]

Kaimami plays an important role in stories, and its role varies across time periods (Hayashida, 1980). In the Nara period and before, *kaimami* plays a role of violation of taboos. That is, man's action of peering at a woman violates the taboo against seeing a woman, which thus often results in punishment. In literary works from the Heian period, especially tales (*monogatari*), Imai (1948/2003) claims that *kaimami* is a technique to create a romantic

---

[28] *Oocho go Jiten* [Dictionary of Heian Lexical Items] (2000)

[29] Different perspectives on *kaimami* have been proposed, such as Yoshikai (2008) and Bargen (2015). Examining *kaimami* scenes with a focus on scenes in *The Tale of Genji*, Yoshikai claims that it is crucial to investigate unconventional forms of *kaimami* and their functions in the story. There are peering scenes in which the observer is a female character and the observed is a male character. In another unconventional form of *kaimami*, Yoshikai (2008) and Bargen (2015) propose, the observed is aware of the existence of the observer and purposefully behaves in a certain way. These types of peering are not conventionally categorized as *kaimami*, but Yoshikai and Bargen claim that unconventional forms of *kaimami* also should be examined to further the interpretation of the text.

relationship between a man and a woman. In the Heian period, Mitani (1994) states, after she attained womanhood, a woman of high rank should not be seen by a man unless he was her father or husband. Women rarely showed their faces even to women unless they lived in the same living sphere. This convention also applies in tales from the Heian period. Though *kaimami*, as Hayashida (1980) and Yoshikai (2008) say, does not always lead to romance or marriage in tales from the Heian period, the concept proposed in Imai (1948/ 2003) has been widely accepted. In addition, Mitani (1994) and Yoshikai (2008) propose another important role of *kaimami*: incorporation of the perceptions of characters (senses of vision, hearing, and smell), feelings and evaluations of characters. That is, *kaimami* has a different function in tales, which contributes to multiple aspects of the story across different time periods. Also, as the previous studies, such as Mitani (1994) and Jinno (2004) discuss, *kaimami* scenes are good place to investigate sentences of represented internal states.

3.5.2. *Kaimami* scenes in *The Tale of Genji*

The Tale of Genji contains multiple *kaimami* scenes. Analyzing the text of *kaimami* scenes in the tale, previous studies of textual analysis on the tale, such as Mitani (1989), Takahashi (1991), and Jinno (2004), have identified sentences of narration that read as if the reader were experiencing the scene or the feeling on his own, mainly by examining the use of honorifics. The *kaimami* scenes are thus important in investigating SRIS rendered in the tale. Yet, the number of *kaimami* scenes varies across researchers. Imai claims that there are 17 *kaimami* scenes in *The Tale of Genji*, while Shinohara (1973) states that there are 55 *kaimami* scenes and Kawakami (1976) proposes that there are 33 *kaimami* scenes in the tale. This difference stems from varying conditions for a *kaimami*. For example, Kawakami includes peering scenes in which the observer is a female character and the observed is a male character while Imai includes only those scenes in which a male character sees a female character. This dissertation thus investigates *kaimami* scenes in *The Tale of Genji* that meet

the following conditions. First, the observer is a male character and the observed is a female character; second, either the observer or the observed is a main character of the tale. There are 18 chapters in which *kaimami* occurs that meet these conditions. Table 3.23 provided below shows the list of the *kaimami* scenes in the tale, which includes a chapter, the title of the chapter, the observer, and the observed.

Table 3.23. *Kaimami* scenes in *The Tale of Genji*

Chapter	Title	The Observer	The Observed
3	Utsusemi	Genji	Utsusemi, Nokiba no ogi
5	Wakamurasaki	Genji	Wakamurasaki (Mursaki no ue), the nun
6	Suetsumuhana	Genji	ladies-in-waiting
7	Momiji no Ga	Emperor Kiritsubo	Genji, Gen no Naishinosuke
21	Otome	Yugiri	Gosechi no Mai-hime
25	Hotaru	Prince Hotaru	Tamakazura
26	Tokonatsu	To no Chujo	Kumoi no Karr, Omi no Kimi
28	Nowaki	Yugiri	Murasaki no Ue, Genji
		Yugiri	Tamakazura, Genji
		Yugiri	Princess Akashi
34	Wakana: jo	Kashiwagi	The Third Princess
40	Minori	Yugiri	Murasaki no Ue
44	Takekawa	Kurodo no Shosho	Tamakazura's daughters
45	Hashihime	Kaoru	The Uji Princesses
46	Shi ga Moto	Kaoru	The Uji Princesses
49	Yadorigi	Kaoru	Ukifune
50	Azumaya	Prince Niou	Ukifune
51	Ukifune	Prince Niou	Ukifune
52	Kagero	Kaoru	The First Princess
53	Tenarai	Chujo	Ukifune

3.5.3. Concordance Plots of *The Tale of Genji*

The current section provides concordance plots focusing on the *kaimami* scenes in *The Tale of Genji* considering the targeted lexical items identified in the section 3.3. (LLR). According to Sinclair (1991), “concordance is a collection of the occurrences of a word-form,

each in its own textual environment. In its simplest form, it is an index. Each word-form is indexed, and a reference is given to the place of each occurrence in a text." (p. 32). In other words, a comprehensive list of a targeted word or pattern is retrieved from the corpus data by conducting a concordance search. The output can be sorted out, for example, by collocation, which facilitates further investigation and leads to a qualitative analysis. Also, the concordance data can be visually presented, which is called a concordance plot. A concordance plot visually shows a distribution of the targeted word or pattern in the text of the corpus.

In *kaimami* scenes, lexical items associated with subjectivity tend to appear in order to render various perceptions as well as the consciousness of a character in the text. *Kaimami* scenes describe not only the action of peering but also the person being peered upon. A characteristic of *kaimami* scenes, Mitani (1994) states, is describing the appearance or attire of the person being peered upon in detail. This detailed description is often rendered through the perception of the character who performs the peering (Mitani, 1994). Visual perception is not the only perception that is rendered in *kaimami* scenes (Mitani, 1994; Yoshikai, 2013). That is, other perceptions, such as auditory perception and olfactory perception, or the consciousness of the observing character, may be rendered in *kaimami* scenes. Describing perceptions and consciousness of a character often require linguistic elements of subjectivity. For instance, the auxiliary verbs *nari* (to indicate speculation based on auditory information) and *meri* (to indicate speculation based on visual information) describe how the observed is seen or heard by a character who is engaging in the peering. Also, the particle for exclamation *kana* is utilized to express a character's emotion, which is stimulated through *kaimami*.

The current concordance search investigates a specific linguistic structure rendered in sentences of narration, in which targeted lexical items are used. Sentences of represented internal states, Mitani (1994) claims, are not cited in sentences of narration though the sentence reads that it directly renders the character's subjectivity, such as thoughts, feelings,

and emotions.[30] Moreover, sentences of represented internal states in the text of *The Tale of Genji*, Mitani demonstrates, become sentences of inner speech of a character if the sentence is followed by a citational marker, as well as an honorific predicate (1994, p. 60). In classical Japanese, cited sentences are followed by a citational marker, such as *to* and *nado*, but not necessarily by a verb, such as *iu* (to say) and *omou* (to think) (Matsushita, 2006). That is, the citational marker indicates the sentence being cited on its own. In other words, the lack of a citational marker contributes to turning the sentence into a sentence that can be read as if the reader were experiencing the scene or the feelings of a character directly. This concordance search investigates where a targeted lexical item (i.e., the auxiliary verb of speculation, the particle for exclamation, and the interjection) is rendered without a citational marker in sentences of narration in the tale. The word order of these lexical items, however, is not considered in the current concordance plots since a citational marker may occur within the sentence in classical Japanese. In addition, the meaning of each lexical item is not considered. For example, the auxiliary verb *beshi* has multiple meanings, such as speculation, evidentiality, and schedule. A sentence that contains *beshi* and does not contain a citational marker therefore may be included even if *beshi* in this sentence indicates schedule, not speculation. Though fine details may be refined, the concordance plots provide a general distribution of the targeted linguistic structure in the sentences of narration in *The Tale of Genji*.

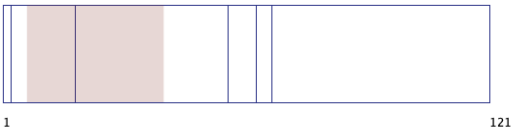
Provided below are the concordance plots for the 18 chapters that contain *kaimami* scenes in *The Tale of Genji*.[31] A vertical line indicates the existence of a sentence in the text type of narration that contains the targeted lexical elements of subjectivity introduced in the section 3.3.7. (i.e., the auxiliary verbs of speculation, the particles for exclamation, and the

[30] Matsushita (2006) states that the character's thoughts and emotions as well as conversations and letters are often quoted in *The Tale of Genji*.

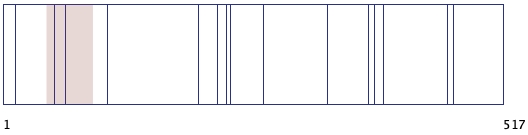
[31] The concordance plots for all the chapters of *The Tale of Genji* are provided in the appendix.

interjections); the number at the bottom of the left hand indicates the first sentence and the number at the bottom of the right hand indicates the last sentences in the chapter. *Kaimami* scenes are highlighted in the concordance plots. For example, in the chapter of *Utsusemi*, there are five 'subjective elements' indicated by five vertical lines, but there is only one that appears in the *kaimami* scene indicated by colored space in the diagram.

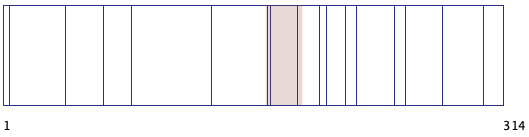
Chapter 3: Utsusemi



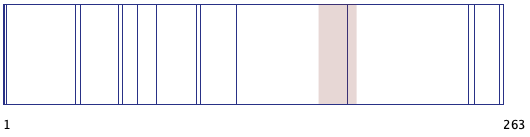
Chapter 5: Wakamurasaki



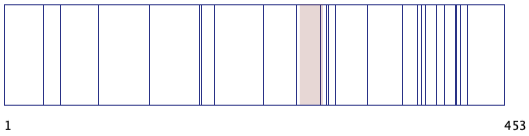
Chapter 6: Suetsumuhana



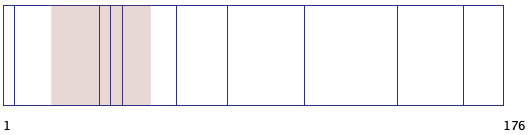
Chapter 7: Momiji no Ga



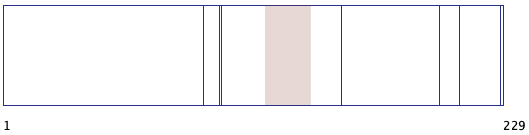
Chapter 21: Otome



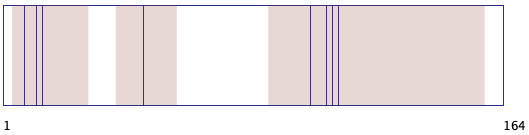
Chapter 25: Hotaru



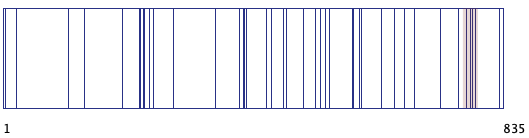
Chapter 26: Tokonatsu



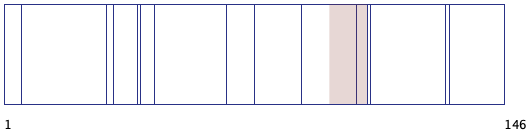
Chapter 28: Nowaki



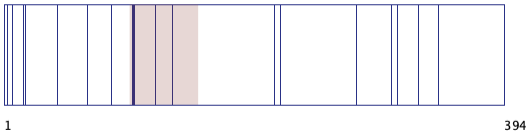
Chapter 34: Wakana: jo



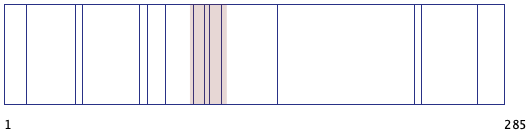
Chapter 40: Minori



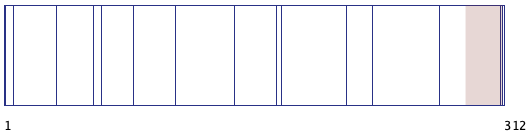
Chapter 44: Takekawa



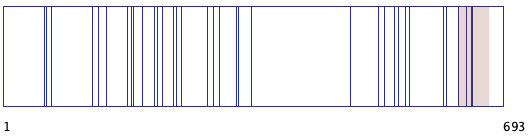
Chapter 45: Hashihime



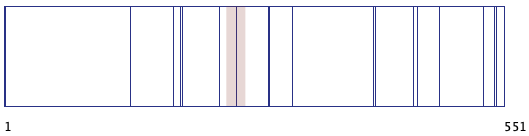
Chapter 46: Shi ga Moto



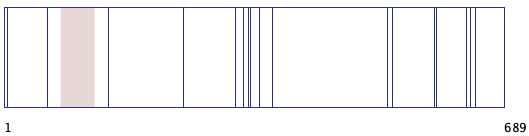
Chapter 49: Yadorigi



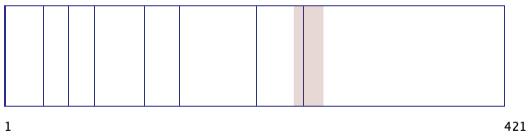
Chapter 50: Azumaya



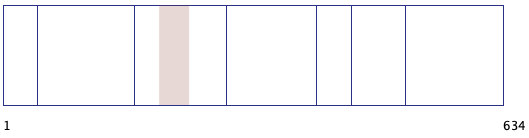
Chapter 51: Ukifune



Chapter 52: Kagero



Chapter 53: Tenarai



As seen above, though some *kaimami* scenes do not contain any sentences of the targeted linguistic structure (Chapter 26 *Tokonatsu*, Chapter 51 *Ukifune* and Chapter 53 *Tenarai*), the concordance plots confirm that out of 20 *kaimami* scenes, most *kaimami* scenes contain sentences of the targeted linguistic structure. That is, lexical items associated with subjectivity appear in most *kaimami* scenes of *The Tale of Genji*. As I mentioned earlier, this frequent occurrence of subjective lexical items in *kaimami* scenes indicates that a character's peering is rendered through his perceptions. Essentially, the person being observed is depicted from the perspective of the person doing the peering, rather than a neutral description. Yet, the targeted structure also occurs in the text outside of the *kaimami* scenes. The occurrence of the targeted linguistic structure outside of the *kaimami* scenes may be due to polysemy of some of the targeted lexical items. For example, without a citational marker, the auxiliary verb *beshi*, as mentioned earlier, is utilized in sentences of narration to indicate schedule while the auxiliary verb *mu* appears in sentences of narration to indicate circumlocution. Also, there may be cases in which sentences of represented internal states may occur outside of *kaimami* scenes. Sentences of represented internal states outside of the *kaimami* scenes may be related to the discussion of the narrator's commentary (*soshiji*), a category of the text classification of *The Tale of Genji* as mentioned earlier. This is another

major research topic in textual analysis of *The Tale of Genji*. However, due to the scope of this dissertation, the current research focuses on the *kaimami* scenes and further analyzes sentences of represented internal states and its textual context.

3.6. Summary

In this chapter, I conducted quantitative analyses (log likelihood ratio and concordance) to determine linguistic elements that contribute to creating sentences of narration in classical Japanese that read as if the reader were experiencing the scene or the feelings on his own. First, the LLR analysis identified word distribution in the three different text types (narration, conversation, short poem) across the 16 literary works from the Heian period. This statistical analysis demonstrated that each text type has its own linguistic preference that is suitable to its function. Lexical items associated with subjectivity (e.g., auxiliary verbs for speculation, exclamatory interjections and particles) which contribute to creating sentences of represented internal states in literary texts (Banfield, 1982; Kuroda, 1973/ 2014), are classified as the highest frequency words for conversation as well as the lowest frequency words for narration. The distribution of these lexical items clearly shows that linguistic elements associated with subjectivity are preferred in sentences of conversation. I then conducted a concordance search for the lexical items associated with subjectivity in the text of *The Tale of Genji*, focusing on sentences of narration. This concordance search was conducted throughout the 54 chapters of the tale, with a specific focus on the *kaimami* scenes in the tale, since *kaimami* scenes were a major focus for investigation of sentences of represented internal states in the previous studies of textual analysis on *The Tale of Genji.* In addition, since omission of citation is a characteristic of sentences of represented internal states, the current concordance search investigated a targeted lexical element utilized in the specific linguistic structure: a targeted lexical element appearing without being cited in narration. The concordance plots demonstrated that a

targeted lexical element is rendered without citation in sentences of narration in *kaimami* scenes in the tale. In conclusion, the current chapter through statistical analysis determined the linguistic elements and confirmed which part of the text of *The Tale of Genji* to further investigate in a qualitative analysis, which leads to the chapter 4.

CHAPTER 4

QUALITATIVE ANALYSIS ON KAIMAMI IN THE TALE OF GENJI

The current chapter investigates sentences of represented internal states (SRIS) rendered in *kaimami* scenes in *The Tale of Genji*. This type of sentence has been intensively investigated in previous textual analyses of the tale, such as Mitani (1994), Takahashi (1991), and Jinno (2004). The previous research mainly focuses on categorizing the text as well as identifying whose voice or perspective is rendered in the text, mostly by analyzing the use of honorifics. Focusing on the "targeted lexical items" (i.e., auxiliary verbs of speculation, particles for exclamation, and interjections), which are found in the quantitative analysis in Chapter 3, the current research examines the way in which these lexical elements are used in SRIS. As explained in the previous chapter (section 3.4), these linguistic elements are associated with subjectivity. This is a common linguistic quality that contributes to creating this phenomenon in literary texts across different languages (cf. Kuroda, 1973/ 2014; Banfield, 1982). Investigating the text of *The Tale of Genji*, this chapter demonstrates how these linguistic elements are also utilized in SRIS in classical Japanese literature.

SRIS in this dissertation are defined as sentences that are rendered in narration by using a "targeted lexical item" without being marked as a citation. The concordance plot research in Chapter 3 identified instances of this type of sentence in *kaimami* scenes in *The Tale of Genji*. The current chapter will examine SRIS appearing in *kaimami* scenes from a qualitative perspective. Out of 20 *kaimami* scenes mentioned in Chapter 3, the current qualitative analysis examines six *kaimami* scenes from five different chapters: the 3rd chapter (*Utsusemi*), the 5th chapter (*Wakamurasaki*), the 28th chapter (*Nowaki*), the 34th chapter (*Wakana: jo*), and the 46th chapter (*Shiigamoto*). These are most famous *kaimami* scenes in *The Tale of Genji* and play important roles in the story.

In the following sections, I will first present the data that is used for the current

qualitative analysis, and briefly discuss issues on the textual tradition of *The Tale of Genji*. In the analysis section, I explain the scope of the current qualitative analysis as compared to the previous studies before presenting my qualitative analysis. The qualitative analysis in this chapter is categorized according to the function of "targeted lexical items" as well as the function of sentences of represented internal states in context.

4.1. Data

The *kaimami* scenes I analyze in this chapter come from the *Shinpen Nihon koten bungaku zenshū*, which contains a selection of important works of classical Japanese literature written from the Nara period (710-794) up until the Edo period (1603-1868). The text of this annotated edition of the tale is one of the most popular among those widely used or referred to for research, textbooks, and translations. However, there has been a major discussion on the textual tradition of *The Tale of Genji* regarding annotated editions and manuscript copies (Kato, 2009). No manuscript copies written in the Heian period (794-1192) have survived. In other words, we no longer have access to any manuscripts written in the era when the tale was written. Yet, Kato states that approximately 150 different manuscripts of the tale have survived from different time periods, and the oldest extant manuscript copy was written in the Kamakura period (1192-1333) by Fujiwara no Teika, which is known as the *Teika-bon*. Those manuscripts written in the Kamakura period and the Northern and Southern Court (*Nambokucho)* period (1336-1392) are, however, all incomplete texts including the *Teika-bon*. That is, there remains no surviving manuscript from those periods that consists of *The Tale of Genji* in its entirety.

Texts of the annotated editions of *The Tale of Genji*, such as the *Shinpen Nihon koten bungaku zenshū* and the *Shin Nihon koten bungaku taikei*, are composed based on different manuscripts. Among approximately 150 extant manuscripts, three manuscript copies have mainly been used as the base texts to compose annotated editions of *The Tale of Genji* in the

postwar period: the *Teika-bon*, the *Myōyū-bon*, and the *Ōshima-bon*. The *Teika-bon*, which was prepared in the early Kamakura period (1192-1333), has four surviving chapters, while the *Myōyū-bon*, which was written in the end of Muromachi period (1338-1573), has 53 surviving chapters, and the *Ōshima-bon*, which was written in the end of the Muromachi period, also has 53 surviving chapters (Kato, 2009). For example, the text of the *Shin Nihon koten bungaku taikei* is composed in reference to the *Ōshima-bon* for most chapters and the *Myōyū-bon* for one chapter while the text of the *Shinpen Nihon koten bungaku zenshū* is composed in reference to four different manuscripts including the oldest manuscript copy the *Teika-bon*. In each case, the choice of the base manuscripts has been made according to the editor's policy. Kato also discusses another issue on the textual tradition of *The Tale of Genji*, which is the selection of these three major manuscripts used to compose the annotated editions of the tale. The *Teika-bon* has been considered to be the primary manuscript because Fujiwara no Teika, who is a major figure in Japanese poetry as well as a scholar of the late Heian period to the early Kamakura period, took part in its composition. The *Myōyū-bon* and the *Ōshima-bon* are, on the other hand, also considered to be important since both manuscripts are thought to inherit the tradition of the *Teika-bon*. However, as Kato claims, it is necessary to examine the validity of these claims on the selection of manuscripts as well as to reconsider whether these three manuscripts are more appropriate to use for the base text over other extant manuscript copies.

These issues on the textual tradition of *The Tale of Genji* have still been under discussion and have not come to a conclusion yet. However, the current qualitative analysis uses the text of the *Shinpen Nihon koten bungaku zenshū* to be consistent with the text of the Corpus of Historical Japanese of the Heian period, which is based on this text, and which forms the basis of the quantitative analysis in Chapter 3. Yet, I also refer to three other major annotated editions: the *Nihon koten bungaku zenshū*, the *Nihon koten bungaku taikei*, and the *Shin Nihon koten bungaku taikei*.

4.2. SRIS in *Kaimami* Scenes in *The Tale of Genji*

The current research investigates SRIS that are rendered in *kaimami* scenes in *The Tale of Genji*. SRIS are sentences of direct representation of a character's consciousness that is read as if the reader were experiencing the character's feelings and emotions on his own. That is, the current qualitative analysis focuses on whose consciousness (thoughts, feelings, and emotions) is rendered in sentences of narration as well as how this type of sentence functions in context.

The current qualitative analysis employs the framework of the non-communicational model, which is introduced in Chapter 2. The framework of the non-communicational model postulates that there is no narrator in a story unless the narrator is a character (Kuroda, 1975/ 2014). In other words, this framework does not assume the omniscient narrator in analysis, who in the framework of the communicational model, is considered to unify the text, provide voice throughout a story and function as the locus of meaning. As discussed in Chapter 2, the assumption of the existence of the omniscient narrator dictates not only the view of the unity of the text but also interpretation of the text. Not imposing the existence of an omniscient narrator in analysis, the framework of the non-communicational model thus makes it possible to examine and interpret the text based on how the language is rendered in the text.

This qualitative analysis, therefore, investigates the "targeted lexical items" (auxiliary verbs of speculation, particles for exclamation and interjections) rendered in SRIS in *kaimami* scenes. These linguistic elements, as discussed in Chapter 3, are all associated with subjectivity. Subjectivity, such as internal state, mental process, and sensation, "is only accessible by the person, who is experiencing or processing them" (Iwasaki, 1993, p.13). In other words, other than the person who is experiencing internal states or processing mental process, no one can directly access subjectivity or epistemic knowledge and be certain about them. This difference in "the accessibility of information" (Iwasaki, 1993) is linguistically

rendered. An example of this is in the morphological differences in modern Japanese between *samishii* 'sad' for the 1st person and *samishi-garu* 'to look sad' for the 2nd/ 3rd person as discussed in Chapter 2 (cf. Kuroda, 1973/ 2014; Iwasaki, 1993). This linguistic difference in information accessibility also applies to SRIS, in which the subjective linguistic elements are used. In modern Japanese, if these subjective linguistic elements (e.g., *samishii* 'sad') are modalized (e.g., *samishi-garu* 'to look sad'), it makes the sentence read as addressed by someone other than the 1st person who is experiencing the sensation. However, in SRIS of modern Japanese, these subjective linguistic elements are not modalized; the absence of modalization makes the sentence read in the first person (Kuroda, 1973/ 2014). In classical Japanese, on the other hand, if a citational marker (e.g., *to* and *tote*) and the following verb phrase including an honorific follow SRIS, it makes the sentence read as the inner speech properly cited by someone other than the 1st person (Mitani, 1994). However, in SRIS of classical Japanese, a citational marker and the following verb phrase are absent; the absence of a citational marker and a citational verb phrase makes the sentence read in the 1st person.[32] Analysis of these linguistic elements associated with subjectivity thus identifies to whom these linguistic elements are attributed, or whose consciousness is realized.

The scope of this research is smaller as compared to the previous research of the textual analysis on *The Tale of Genji*, which investigates not only a character's consciousness but also a character's perspective rendered in sentences of narration. That is, the previous studies examine the voice as well as the perspective. Perspective may also be rendered as the direct representation of a character's "non-verbalized perception" (Prince, 1987/ 2003). However, investigation of a character's perspective requires analysis of linguistic elements

[32] Some examples that Mitani (1994) listed as SRIS do not contain any subjective linguistic elements that are identified in the current dissertation.

other than the current "targeted lexical items," such as deixis (Short, 1996).[33] This dissertation, however, targets, as has been mentioned repeatedly, the following linguistic elements: auxiliary verbs for speculation, particles for exclamation and interjection. Therefore, the current analysis focuses on the consciousness of characters directly rendered in sentences of narration.[34] In the analysis of a *kaimami* scene in the 3rd chapter *Utsusemi*, I examine how a character's speculation based on auditory information is rendered in direct representation while in the analysis of a *kaimami* scene in the 5th chapter *Wakamurasaki*, I analyze how a character's speculation based on visual information is rendered in SRIS. The investigation of a *kaimami* scene from the 46th chapter *Shiigamoto* examines how visual and auditory information is rendered in the direct representation of a character's speculation; the examination of *kaimami* scenes from the 28th chapter *Nowaki* and the 34th chapter *Wakana: jo* focuses on the function of direct representation of a character's thought and consciousness in context. Each sentence provided in the excerpts below is accompanied by three-line gloss: romanization in English, morpheme-by-morpheme gloss, and English translation by Seidensticker. The abbreviation of each gloss is listed in the list of the abbreviation.

[33] Deictic expressions are one of the linguistic elements that Iwasaki (2013) claims are associated with subjectivity. Iwasaki states that deictic expressions, such as demonstratives and movement and transaction expressions, "encode information relevant to the speaker's here-and-now orientation" (p.290). In other words, deictic expressions situate the speaker at the ultimate focal point, which defines other spatial locations in the context. Therefore, investigating these linguistic elements reveal whose perspective is employed to portray a particular scene as well as how a fictional world is portrayed in the text.

[34] In most previous studies of *The Tale of Genji* (e.g., Mitani, 1994; Takahashi, 1991; Jinno 2004), the question of "who sees?" is amalgamated with the question of "who speaks?". The amalgamation of these two questions is considered to be a major problem in narratology; it is imperative to distinguish between "who sees?" and "who speaks?" (Gunette, 1980; Rimmon-Keenan, 1983, Bal, 1985; Yamaoka, 2000). However, scholars of textual analysis on *The Tale of Genji*, such as Takahashi (1991) and Jinno (2004), claim that it is impossible to draw a distinction between these two matters in textual analysis on the tale.

4.2.1. Direct Representation of Speculation based on Visual Information Rendered in Narration: The *Kaimami* Scene in the 5th Chapter *Wakamurasaki*

The *kaimami* scene in the fifth chapter, which is one of the most famous *kaimami* scenes in *The Tale of Genji*, has been extensively investigated for SRIS in the previous studies of textual analysis on *The Tale of Genji*. In this chapter, Genji performs *kaimami* when he visits a holy monk in a mountain to receive medical treatment for an illness. One evening, during his stay, Genji takes a walk in the neighborhood with his retainer Koremitsu and happens to find a girl known as Wakamurasaki, who will go on to become his beloved wife in the later part of the tale, and a nun (Wakamurasaki's grandmother) through the action of peering. In this *kaimami*, the observer is Genji; the observed are Wakamurasaki and the nun. SRIS occur soon after Genji starts the action of peering (*kaimami)*.

Excerpt 1: The *kaimami* in the 5th chapter *Wakamurasaki*

Line (1)

人々	は	帰し	たまひ	て
hitobito	wa	kaeshi	tamai	te
N: people	P (bound)	V: to send back	V (respect)	P (conjunctive)

惟光	朝臣	と	のぞき	たまへ	ば
koremitsu	ason	to	nozoki	tamae	ba
N (name)	N (title)	P (case)	V: to peer	V (respect)	P (conjunctive)

Sending back everyone except Koremitsu, he took up a position at the fence

Line (2)

ただ	こ	の	西面	に	しも
tada	ko	no	nishiomote	ni	shimo
ADV: close	PN	P (case)	N: a room facing the west	P (case)	P (adverbial)

持仏	すゑ	たてまつり	て	行ふ
jibutsu	sue	tatematsurite	te	okonau
N: Buddha statue	V: to place	V (humbleness)	P (conjunctive)	V: to practice

尼	なり	けり
ama	nari	keri
N: nun	AUX	AUX

In the west room sat a nun who had a holy image before her

Line (3)

簾	すこし	上げ	て
sudare	sukoshi	age	te
N: blind	ADV: slightly	V: to raise	P (conjunctive)

花	奉る	**めり**
hana	tatematsuru	**meri**
N: flower	V (humbleness)	**AUX: supposition based on visual evidence**

The blinds were slightly raised and she seemed to be offering flowers

Lines (1) (2) and (3) are all rendered in sentences of narration. Bringing only his retainer Koremitsu, Genji comes up to a fence and looks at the house behind it (Line 1). In Line (1), the beginning of *kaimami* is indicated by the verb *nozoki* (*nozoku*: to peer). What first catches Genji's eyes is the nun, witting with a Buddhist statue in front of her (Line 2). Having the blind slightly rolled up, the nun looks like she is offering flowers to the Buddha statue (Line 3), which is rendered in the form of SRIS by using the auxiliary verb *meri* (supposition based on visual evidence) without being followed by a citational marker. The use of the auxiliary verb *meri* indicates that Genji's conjecture is based on the visual information available to him. In other words, Genji cannot see the nun's action clearly probably due to the blinds;

therefore, he makes a guess.

Another instance of SRIS occurs when Genji is observing a lady-in-waiting, who is taking care of Wakamurasaki.

Excerpt 2: The *kaimami* in the 5th chapter *Wakamurasaki*

Line (4)

髪	ゆるるかに	いと	長く
kami	yururukani	ito	nagaku
N: hair	ADJD: rich	ADV: very	ADJ: long

めやすき	人	な	**めり**
meyasuki	hito	na	**meri**
ADJ: pleasant	N: person	AUX	**AUX: supposition based on visual evidence**

… said a rather handsome woman with rich hair

Line (5)

少納言	の	乳母	と	ぞ
shonagon	no	menoto	to	zo
N (title)	P (case)	N: nurse	P (case)	P (bound)

人	言ふ	**める**	は
hito	iu	**meru**	wa
N: person	V: tosay	**AUX: supposition based on visual evidence**	P (bound)

(a rather handsome woman) who seemed to be called Shonagon

Line (6)

こ	の	子	の
ko	no	ko	no
PN	P (case)	N: child	P (case)

後見	なる	**べし**
ushiromi	naru	**beshi**
N: guardian	AUX	**AUX: conjecture with confidence**

[a rather handsome woman…who] was apparently the girl's nurse

Lines (4) (5) and (6) are rendered in the form of direct representation by using the auxiliary verbs *meri* (supposition based on visual evidence) and *beshi* (conjecture with confidence) without any other elements, such as a citational marker, that attribute the subjectivity to someone other than the 1st person. Genji sees the lady-in-waiting who has very long hair, and the use of the auxiliary verb *meri* indicates that he makes a judgment on her appearance based on the visual information available to him: she looks attractive to him (Line 4). Line (4) is immediately followed by Lines (5) and (6). It seems like she is called Shonagon (Line 5), and she must be a nurse of this little girl (Line 6). The auxiliary verb *meri* used in Line 5 indicates that this conjecture on the name of this lady-in-waiting is made mainly through his action of peering. In Line 6, the auxiliary verb *beshi* is utilized to suggest his speculation on her role as a nurse of Wakamurasaki is processed through his logical thinking. As discussed above, without the assumption of the existence of the omniscient narrator, there is no linguistic trace to indicate its existence in these instances. Rather, the auxiliary verbs utilized in Lines (4), (5), and (6) all indicate Genji's subjective judgment considering that he is the observer in this *kaimami* scene.

These instances of SRIS, as seen above, demonstrate that SRIS provide two effects in the text. First, SRIS add dynamicity to the text. The effect of vivid dynamicity is seen in

Line (5) and Line (6). The both examples are direct representation of Genji's thought (conjecture on the appearance of the lady-in-waiting called Shonagon as well as her role as a nurse). Instead of citing his thought, these examples directly convey Genji's conjecture by utilizing the auxiliary verbs *meri* (supposition based on visual evidence) and *beshi* (conjecture with confidence). These linguistic elements, as discussed in the previous chapter, are associated with subjectivity. Since in this *kaimami* scene, Genji is the observer as well as the one who makes a conjecture or judgment, these lexical items associated with subjectivity are thus attributed to him. That is, without any outside interference from extra information or citations, the sentence reads as if the reader were experiencing the scene or the feeling of Genji on his own, which brings in vividness to the scene. Secondly, SRIS may add depth to the description of a scene by providing a limited amount of information, instead of providing the full information of a scene. This effect is seen in Line (3), in which the auxiliary verb *meri* is rendered in SRIS. In this scene, Genji peers upon the nun, who seems to be offering flowers. The auxiliary verb *meri* in Line (3) is used to describe the nun's action through Genji's visual perception; he cannot clearly see her maybe due to the blinds that block his sight. By providing limited visual information given to the reader, it indicates unseen parts of the scene. This indication of unseen parts in the scene adds depth to its visual description.

4.2.2. Direct Representation of Speculation based on Auditory Information Rendered in Narration: The *Kaimami* Scene in the 3rd Chapter *Utsusemi*

In the *Utsusemi* chapter, Genji peers at two women when he visits the residence of the governor of Kii. During his visit, through the action of *kaimami*, he sees Utsusemi (the governor's second wife) and Nokiba no Ogi (the governor's daughter) playing a game of *go*. Though Genji has already met Utsusemi during his previous visit to the residence, this is the first time for him to see her. Genji is waiting outside of the room for a chance to meet with Utsusemi while the two women play a game of *go*. An instance of SRIS occurs when Genji

quits peering and engages in a conversation with Kogimi (Utsusemi's younger brother), who just came back from the room, where the two women are playing a game of *go*, and tells him of his plan to let Genji go inside when Nokiba no Ogi leaves.

Excerpt 3: The *kaimami* in the 3rd chapter *Utsusemi*

Line (1)

碁	打ち	はて	つる	に	や	あら	**む**
go	uchi	hate	tsuru	ni	ya	ara	**mu**
N	V	V	AUX	AUX	P (bound)	V	**AUX: conjecture**

The game of *go* was apparently over

Line (2)

うち	そよめく	ここち	し	て
uchi	soyomeku	kokochi	shi	te
PREF	V: to rustle	N: feeling	V: to do	P (conjunctive)

人々	あかるる	けはひ	など	す	**なり**
hitobito	akaruru	kahai	nado	su	**nari**
N: people	V: to withdraw	N: atomosphere	P (adverbial)	V: to do	**AUX: auditory speculation**

There was a stir inside, and a sound as of withdrawing

In this *kaimami* scene, the observer is Genji and the observed are Utsusemi and Nokiba no Ogi. Line (1) and Line (2) are rendered in the form of direct representation of thought, by using three lexical items associated with subjectivity: the particle *ya* (doubt/ question), the auxiliary verb *mu* (conjecture), the auxiliary verb *nari* (assumption based on sound). Genji is not certain, yet it feels like the play of *go* is now over (Line 1). His conjecture is rendered in

the form of SRIS by using the particle *ya*, which indicates uncertainty, and the auxiliary verb *mu*, which indicates speculation, without being cited. Genji is not currently observing inside, and he has to guess about what the situation was inside of the room. This conjecture is followed by his speculation based on auditory information. Genji hears clothes rustling from inside, which indicates that people are leaving the room (Line 2). This speculation based on auditory information given to Genji is also rendered in the form of SRIS by utilizing the auxiliary verb *nari* (assumption based on sound). In addition to *nari*, the noun *kehai* 'appearance,' which indicates a feeling acquired by smelling, hearing or touching, is also utilized to show that Genji is making a conjecture based on auditory information. [35]

These instances are followed by another instance of SRIS when Genji hears people inside preparing to go to sleep.

Excerpt 4: The *kaimami* in the 3rd chapter *Utsusemi*

Line (3)

「	若	君	は	いづく	に	おはします	なら	**む**
	waka	gimi	wa	izuku	ni	owashimasu	nara	**mu**
	N: young boy	SUFF	P (bound)	PN	P (case)	V: to be (respect)	AUX	**AUX: conjecture**

こ	の	御	格子	は	鎖し	て	ん	」	とて
ko	no	mi	kooshi	wa	sashi	te	mu		tote
PN	P (case)	PREF	N: door	P (bound)	V: to close	AUX	AUX		P (case)

"Where will that boy have gone?"… "Let's get the place close up."

[35] The noun *kehai* is a lexical item of highest frequency word for narration as well as of lowest frequency word for conversation in the LLR analysis (Chapter 3).

Line (4)

鳴らす	**なり**
narasu	**nari**
V: to make noise	**AUX: auditory speculation**

Now there was a banging of shutters.

Here, Genji is still observing the inside. Immediately after the play of *go* is over, a lady-in-waiting inside says that she will close the door, which is rendered in the sentence of conversation followed by the citational marker *tote* (Line 3). Then, Genji hears the door being closed (Line 4). The auxiliary verb *nari* utilized in Line (4) indicates that based on his auditory information (the lady-in-waiting's statement and the banging of shutters) Genji assumes that they closed the door.[36]

In the *kaimami* scene from the *Utsusemi* chapter, SRIS add dynamicity to the scene as in the *Wakamurasaki* chapter. While in the *Wakamurasaki* chapter, the effect of vividness is added through rendering visual information in the form of direct representation, in the *Utsusemi* chapter, the effect of vividness is added through rendering auditory information in the form of direct representation. In the *Utsusemi* chapter, Genji is the observer and the lexical items that are associated with subjectivity discussed above are all attributed to him. Due to the absence of any interference from extra information or citation, Genji's subjectivity is directly rendered in the text. In this scene, waiting outside of the room, Genji is attentive to the situation inside so that he will not miss a chance to meet with Utsusemi. Therefore, the information from the inside of the room plays an important role here. Describing Genji's conjecture about the situation inside of the room using SRIS brings life to the scene by directly providing the auditory information available to Genji, instead of a mere description

[36] Oda (2015) explains that the auxiliary verb *nari* is used for a situation where one does not have access to visual information and has to make a conjecture solely based on his auditory information.

of the situation. This way, the reader feels as if he were experiencing on his own here-and-now atmosphere as well as the suspenseful, tense feelings that Genji experiences in the scene.

4.2.3. Direct Representation of Speculation based on Both Visual and Auditory Information Rendered in Narration: The *Kaimami* Scene in the 46th Chapter *Shiigamoto*

The *Shiigamoto* chapter is a story about Genji's offspring after his passing. In this chapter, Kaoru (known as the youngest son of Genji, though he is actually an illegitimate child between the Third Princess, a wife of Genji, and Kashiwagi, a friend of Genji's son) visits the Uji villa, which is located outside of the city. Kaoru goes to Uji to pay a visit to the Uji Princesses, who are under his guardianship due to their father's will. During his visit, Kaoru happens to get a glimpse of the princesses. In this *kaimami* scene, the observer is Kaoru and the observed are the Uji Princesses.

Excerpt 5: The *kaimami* in the 46th chapter *Shiigamoto*

Line (1)

外	に	立て	たる	屏風	を
to	ni	tate	taru	byoobu	o
N: outside	P (case)	V: to place	AUX	N: screen	P (case)

ひきやり	て	見	たまふ
hikiyari	te	mi	tamau
V: to remove	P (conjunctive)	V: to see	V (respect)

In great exccitement, he pulled aside the screen before the partition

Line (2)

ここ	もと	に	几帳	を	そへ	立て	たる
koko	moto	ni	kichoo	o	soe	tate	taru
PN	N: vicinity	P (case)	N: screen	P (case)	V: to add	V: to place	AUX

あな	口惜し	と	思ひ	て
ana	kuchioshi	to	omoi	te
INTJ	ADJ: unfortunate	P (case)	V: to think	P (conjunctive)

ひき	帰る	をり	しも
hiki	kaeru	ori	simo
V: to withdraw	V: to return	N: at that time	P (adverbial)

Alas, there was a curtain beyond. But as he drew back

Line (3)

風	の	簾	を	いたう	吹き	上ぐ
kaze	no	sudare	o	itoo	fuki	agu
N: wind	P (case)	N: blind	P (case)	ADV: badly	V: to blow	V: to go beyond

べか	**めれ**	ば
beka	**mere**	ba
AUX: conjecture with confidence	**AUX: supposition based on visual evidence**	P (conjunctive)

… the wind caught the blind at the front veranda

Line (4)

「 あらはに	も	こそ	あれ
arawani	mo	koso	are
ADJV: exposed	P (bound)	P (bound)	V: to be

そ	の	御	几帳	押し	出て	て	こそ 」
so	no	mi	kichoo	oshi	ide	te	koso
PN	P (case)	PREF	N: screen	V: to push	V: put out	P (conjunctive)	P (bound)

Pull them over, hold it down…"The whole world can see us."

Line (5)

と	言ふ	人	あ	**なり**
to	iu	hito	a	**nari**
P (case)	V: to say	N: person	V: to be	**AUX: assumption based on sound**

…said someone…

Upon arriving, Kaoru is having a rest in a room and senses that the princesses, who happen to be in the next room, are returning to their own room. These rooms are separated by a single partition, in which Kaoru has found a hole beforehand. In order to peer in at the princesses in the next room, he pushes aside a screen placed in front of him, behind which the partition is placed, and looks into the small hole in the partition (Line 1). However, there is another screen on the other side of the partition, which prevents Kaoru from seeing the Uji Princesses. Disappointed, he is about to quit peering and return to his seat (Line 2). SRIS are rendered when Kaoru is about to quit peering. Just as he is about to return to his seat, the wind blows up a blind that is placed at the veranda in the princesses' room to block the sight from outside (Line 3). Line 3 describes visual information given to Kaoru, which is rendered in the form of SRIS by utilizing the auxiliary verbs *beshi* (conjecture with confidence) and *meri* (supposition based on visual evidence). These two auxiliary verbs indicate that Kaoru cannot see clearly, yet it seems as if the blind has been blown up by the wind. Immediately after he sees what is going on, a lady-in-waiting tells others to pull over the curtain placed on

the other side of the partition that blocks Kaoru's sight, which is rendered in conversation (Line 4). Kaoru hears a lady-in-waiting say so, which is rendered in the form of direct representation (Line 5). This utterance is auditory information given to Kaoru, which is indicated by the auxiliary verb *nari* (assumption based on sound).

In the *Shiigamoto* chapter, SRIS are utilized to describe Kaoru's *kaimami* of the Uji Princesses, which is achieved due to the visual and auditory information given to Kaoru. That is, Kaoru recognizes a change caused by the wind and then he perceives the information that the curtains will be moved away, which suggests that Kaoru is given the opportunity to peer at the princesses. Rendering these clues through direct representation of Kaoru's consciousness adds dynamicity, which sets the mood for the following *kaimami* of the Uji Princesses performed by Kaoru. In addition, rather than a mere description of the situation, SRIS make it possible for the reader to feel as if he were experiencing the scene on his own.

4.2.4. Direct Representation of the Character's Heart Rendered in Narration

4.2.4.1. The *Kaimami* Scene in the 28th Chapter *Nowaki*

In the *Nowaki* chapter, Yugiri (Genji's son) peers at multiple women at the Rokujo estate, which is Genji's residence. One autumn day, Yugiri pays a visit to the Rokujo palace after a typhoon. During his first visit to the palace, Yugiri happens to see Murasaki no Ue, who is a beloved wife of Genji as well as Yugiri's stepmother. During his second visit, he peers in on Tamakazura, who he thinks is a half-sister though she is actually a cousin, and Princess Akashi, who is his younger sister. SRIS are first utilized when Yugiri peers in on Murasaki no Ue by accident, who is standing at the veranda looking at the flowers in the garden.

Excerpt 6: The *kaimami* in the 28th chapter *Nowaki*

Line (1)

御簾	の	吹き	上げ	らるる	を
misu	no	fuki	age	raruru	o
N: screen	P (case)	V: to blow	V: go beyond	AUX	P (case)

人々	押へ	て
hitobito	osae	te
N: people	V: to hold	P (conjunctive)

[She laughed] as her women fought with the unruly blinds

Line (2)

いかに	し	たる	に	か	あら	**む**
ikani	shi	taru	ni	ka	ara	**mu**
ADV: how	V: to do	AUX	AUX	P (bound)	V: to be	**AUX: conjecture**

...though he was too far away to make out what she said to them

Line (3)

うち	笑ひ	たまへ	る	いと	いみじく	見ゆ
uchi	warai	tamae	ru	ito	imijiku	miyu
PREFF	V: to laugh	V (respect)	AUX	ADV: very	ADJ: beautiful	V: to look

She (Murasaki no ue) laughed…the bloom was more radiant

In this scene, Yugiri is astonished by Murasaki's beauty. Upon arriving at the Rokujo estate, Yugiri looks inside. The partitions in Murasaki no Ue's room are all pulled aside due to the strong wind, which makes it possible for Yugiri to see inside. While he is peering, the wind blows up the blinds and the ladies-in-waiting pull them down (Line 1). Seeing the ladies-in-

waiting struggling with the blinds, Murasaki no Ue smiles beautifully (Line 3), which makes Yugiri wonder what has happened (Line 2). Line (2) uses SRIS—specifically by using the particle *ka* (question marker) and the auxiliary verb *mu* (conjecture)—to render Yugiri's consciousness (Line 2). This sentence does not describe his visual or auditory conjecture but rather what is evoked in Yugiri's mind based on what he sees. His thought process, as it is rendered in this sentence, indicates that he does not have full access to the situation of Murasaki no Ue. Therefore, visual information is the prime information given to Yugiri, which leaves in his mind only the impression of her astonishing beauty. In addition, incorporating Yugiri's heart into the description of this *kaimami*, instead of a mere description of the appearance of Murasaki no Ue, adds vividness to the text and describes how captivated Yugiri is.

Another instance of SRIS in the *Nowaki* chapter occurs during Yugiri's *kaimami* on Tamakazura. On the following day, when Genji pays a visit to Tamakazura bringing Yugiri along with him, Yugiri has a chance to catch a glimpse of her. Upon seeing Tamakazura, Yugiri is again attracted by her beauty and compares her with Murasaki no Ue, whom he has seen the day before.

Excerpt 6: The *kaimami* in the 28th chapter *Nowaki*

Line (4)

八重	山吹	の	咲き	乱れ	たる
yae	yamabuki	no	saki	midare	taru
N: multi-layered	N: yamabuki flower	P (case)	V: to bloom	V: to be in profusion	AUX

盛り	に	露	かかれ	る	夕映え	ぞ
sakari	ni	tsuyu	kakare	ru	yuubae	zo
N: peak	P (case)	N: dew	V: to place	AUX	N: sparkling in the evening	P (bound)

ふと	思ひ	出で	らるる
futo	omoi	ide	raruru
ADV: suddenly	V: to call to mind	V: to put out	AUX

He thought of a rich profusion of yamabuki sparkling with dew in the evening twilight

Line (5)

をり	に	あは	ぬ	よそへ	ども	なれ	ど
ori	ni	awa	nu	yosoe	domo	nare	do
N: season	P (case)	V: to fit	AUX	N: analogy	SUFF	AUX	P (conjunctive)

なほ	うち	おぼゆる	やう	**よ**
nao	uchi	oboyuru	yoo	**yo**
ADV: all the same	PREF	V: to feel	N: appearance	**P (final): exclamation**

The image was of spring and not autumn, of course, but it was the one that came to him all the same

Tamakazura's beauty reminds him of beautiful *yamabuki* flowers blooming in profusion in the evening (Line 4). In Line (4) the auxiliary verb *raruru* (*raru*) is used, which is not an auxiliary verb of speculation yet indicates an action or situation occurring spontaneously without any intention. That is, the use of this auxiliary verb describes Yugiri's spontaneous reaction to seeing Tamakazura's beauty, which is considered to be his internal thought. Line (5) immediately follows Line (4). Since *yamabuki* are spring flowers, such an analogy may not be suitable for the current season (the autumn). Yet, it is of *yamabuk*i flowers that Tamakazura's beauty reminds him (Line 5). Yugiri's thoughts on Tamakazura's beauty are rendered in the form of SRIS by utilizing the particle *yo* (exclamatory) without being followed by a citational marker. Again, instead of a mere description of his visual information, the direct representation of Yugiri's impression adds vividness to the scene as well as to the description of how attracted he is at the same time.

4.2.4.2. The *Kaimami* Scene in the 34th Chapter *Wakana: jo*

Another instance of this type of SRIS is found in a *kaimami* scene in the 34th chapter. *Kaimami* in this chapter occurs at the Rokujo estate by Kashiwagi, a friend of Yugiri. He has longed for a marriage with the Third Princess (Genji's wife), yet his hope was not granted because Emperor Suzaku (the father of the Third Princess) arranged a marriage between the Third Princess and Genji. However, Kashiwagi cannot give up on her and still has feelings for the princess even after her marriage to Genji. One day, there is a gathering at the Rokujo palace organized by Yugiri, and young gentlemen play a ball game called *kemari*. During the game, Yugiri and Kashiwagi take a rest at the stairs, which are rather close to the Third Princess's rooms. While they are resting, a little cat comes running out of the princess's rooms while being pursued by another large cat. While the little cat is trying to run away from the big cat, its leash is accidentally caught in the blinds placed in the princess's room. The blinds are pulled back, and the Third Princess is suddenly revealed to Kashiwagi.

Excerpt 7: The *kaimami* in the 34th chapter *Wakana: jo*

Line (1)

几帳	の	際	すこし	入り	たる	ほど	に
kichoo	no	kiwa	sukoshi	iri	taru	hodo	ni
N: screen	P (case)	N: edge	ADV: slightly	V: to go inside	AUX	N:vicinity	P (case)

袿姿	に	て	立ち	たまへ	る	人	あり
uchikisugata	ni	te	tachi	tamae	ru	hito	ari
N: being in an informal dress	AUX	P (conjunctive)	V: to stand	V (respect)	AUX	N: person	V: to be

A lady in informal dress stood just inside the curtains

Line (2)

階	より	西	の	二	の	間	の
hashi	yori	nishi	no	ni	no	ma	no
N: footstep	P (case)	N: west	P (case)	N: second	P (case)	N: between	P (case)

東	の	そば	なれ	ば
hingashi	no	soba	nare	ba
N: east	P (case)	N: vicinity	AUX	(conjunctive)

紛れどころ	も	なく	あらわに	見入れ	らる
magiredokoro	mo	naku	arawani	miire	raru
N: hidden place	P (bound)	ADJ: absent	ADJC: exposed	V: to look inside	AUX

[A lady in informal dress stood just inside the curtains] beyond the second pillar to the west

Line (3)

紅梅	に	や	あら	**む**
koobai	ni	ya	ara	**mu**
N: the color of the layered clothes (red and lavender)	AUX	P (bound)	V: tobe	**AUX: conjecture**

濃き	薄き	すぎすぎに	あまた	重なり	たる	けぢめ
koki	usuki	sugisugini	amata	kasanari	taru	kejime
ADJ: deep	ADJ: light	ADV: one after another	N: many	V: to layer	AUX	N

はなやかに	草子	の	つま	の
hanayakani	sooshi	no	tsuma	no
ADJV: bright/ colorful	N: book	P (case)	N: edge	P (case)

やう	に	見え	て	桜	の
yoo	ni	mie	te	sakura	no
N: appearance	AUX	V: to look	P (conjunctive)	N: cherry blossom	P (case)

織物	の	細長	なる	**べし**
orimono	no	hosonaga	naru	**beshi**
N: woven fabric	P (case)	N: clothes for noble women	AUX	**AUX: conjecture with confidence**

Her robe seemed to be of red lined with lavender,
and at the sleeves and throat the colors were as bright and varied as a book of paper samples
Her cloak was of white figured satin lined with red.

Inside the rooms, there is a lady in an informal dress standing behind curtains (Line 1). Since the lady is standing at the second pillar from the west, she is fully exposed to Kashiwagi's view (Line 2). Kashiwagi observes her clothes in detail, which we see from the form of SRIS by using the particle *ya* (doubt/ question) and the auxiliary verbs *mu* (conjecture) and *beshi* (conjecture with confidence) (Example 3). Since he has never seen the Third Princess, Kashiwagi does not have a clue of what she looks like. However, in the Heian period, noblewomen wore informal dresses in their own rooms while her ladies-in-waiting wore rather formal dresses to serve their ladies (Yoshikai, 2008). Kashiwagi thus observes the dress that the lady is wearing in great detail, which is distinctive from the rest of the women in the rooms. Instead of a mere description of his visual information, direct representation of Kashiwagi's consciousness describes his observation vividly.

As discussed above, SRIS that are used in the *kaimami* scenes from the *Nowaki* chapter and the *Wakana: jo* chapter demonstrate that direct representation of characters' consciousness, instead of other forms such as a mere description of characters' visual information, not only adds vividness to the scene but also displays the subtlety of characters' consciousness dynamically.

4.3. Summary

In this chapter, I examined SRIS rendered in the *kaimami* scenes of five chapters in *The Tale of Genji*. The current qualitative analysis investigated how a "targeted lexical item"

is rendered in sentences of narration and analyzed the function of this type of sentence in context. The "targeted lexical items" in this chapter are linguistic elements identified in Chapter 3, which are all associated with subjectivity: auxiliary verbs for speculation, particles for exclamation and interjections. In the current qualitative analysis, most SRIS use a targeted auxiliary verb, and some sentences of direct representation utilize a targeted particle for exclamation. However, no SRIS uses targeted interjections for exclamation. Among various targeted auxiliary verbs, *meri*, *nari* and *beshi* are prone to be utilized in the *kaimami* scenes examined in this chapter. The auxiliary verb *meri* indicates speculation based on visual evidence and the auxiliary verb *nari* indicates supposition based on sound evidence while the auxiliary verb indicates conjecture with confidence. Frequent use of *meri* and *nari* in SRIS supports the claim proposed by Yoshikai (2008) and Mitani (1994) that *kaimami* scenes require the presence of not only peering but also eavesdropping. In other words, during performing *kaimami*, characters not only "see" but also "hear."

In the *kaimami* scenes examined in this chapter, SRIS play several important roles in context. without any interference such as citation or a mere description, the perception and/ or consciousness of a character that is conveyed through using linguistic elements associated with subjectivity is directly displayed in the text. This direct representation of a character's subjectivity may add vividness to the text and makes this type of sentence read as if the reader were experiencing the scene or feelings on his own. The sense of vividness is caused by an overlap between the awareness on perception of the character and that of the reader. That is, when the reader reads SRIS, which directly display the character's subjectivity in sentences of narration, he experiences the consciousness, perception or the feelings of the character vicariously through the text. This vicarious experience brings the sense of vividness into the text. In this sense, SRIS are different from a mere description or citation of a character's consciousness or perception. In addition to the effect of vividness, SRIS make it possible to provide in the text visual or auditory information available only to a character in

the scene. In other words, SRIS may not provide the reader with full visual or auditory information on the scene. By providing a limited amount of information that is visually or aurally acquired through a character's perception, SRIS may highlight a piece of information in context, which may play an important role in the story. This can be seen, for example, in Line (2) and Line (4) in the *kaimami* scene of the 3rd chapter *Utsusemi*. The auxiliary verb *nari* (assumption based on sound) is rendered in the form of SRIS and conveys the auditory information available to Genji. Auditory information plays a crucial role in this *kaimami* scene, since it provides information about what is happening inside to Genji, who is waiting outside of the room for a chance to meet with Utsusemi. Therefore, the auditory information rendered in Line 2 (people inside are leaving the room) and in Line 4 (the lady-in-waiting closed the door to go to sleep) is pivotal. Because it limits the amount of information given to the reader, SRIS make it possible to highlight auditory information that would have only been available to Genji. Another important role of rendering a limited amount of information in SRIS is to add depth to the visual or auditory description in the text. Describing a scene by utilizing the auxiliary verb *meri*, for instance, SRIS present the visual information only available to the character who acquires it. The sight may be described to be clearly, vaguely or little seen, which may also indicate unseen parts of the scene. This can be seen in Line (3) in the *kaimami* scene of the 5th chapter *Wakamurasaki*, where Genji peers upon the nun. The nun's action is described through Genji's visual perception by utilizing the auxiliary verb *meri* (supposition based on visual evidence) rendered in SRIS. That is, it limits the visual information given to the reader. By doing so, SRIS also indicate the unseen parts of the description the nun's activities. In other words, by indicating the unseen parts of the scene, SRIS may add depth to its visual description. In conclusion, as examined in this chapter, the sentence of direct representation of a character's consciousness contributes to creating various effects in the text by intertwining different aspects (visual information, auditory information, and even internal thoughts/ feelings) for the reader.

CHAPTER 5

CONCLUSION

5.1. Summary of Findings

Sentences of represented internal states render a character's perceptions, feelings, and emotions in narration without being overtly marked by a citational marker. That is, these sentences do not contain any linguistic elements that explicitly attribute the subjectivity to someone other than a character who is experiencing the consciousness, perception or emotions. It has been intensively examined in the textual analysis of *The Tale of Genji*, starting with the groundbreaking theory proposed by Tamagami, the *Monogatari-ondokuron*, in the 1950s. Despite the large amount of research conducted in the field (e.g., Itoi, 1992; Jinno, 2004; Mitamura, 1996; Mitani, 1994; Takahashi, 1991), previous studies have been mainly limited to qualitative analysis with a heavy focus on the usage of honorifics, which has left other linguistic elements unexamined. Thanks to the development of computational technology, quantitative methods are now available for the textual analysis of classical Japanese. This dissertation has thus employed two different approaches (quantitative analysis and qualitative analysis), focusing on lexical items associated with subjectivity (the auxiliary verb for speculation, the particles for exclamation and interjections), which are important linguistic elements to create sentences of represented internal states (Kuroda, 1973/ 2014; Banfield, 1982).

In Chapter 3, I employed the log likelihood ratio (LLR) on the text of the 16 classical Japanese literary works written in the Heian period including *The Tale of Genji* to perform a quantitative analysis. Identifying the overall word distribution of the 16 literary works across three different text types (narration, conversation, short poem), the LLR analysis has statistically indicated that each text type has its own linguistic preferences that meet its function in the text. The text types of narration and conversation show a

complementary distribution between highest frequency words (HFq) as well as lowest frequency words (LFq). Sentences of narration prefer linguistic elements that are used as conjunction particles, such as *te*, *ni*, and *ba*, and that are categorized as tense or aspect, such as *ri*, *tari* and *keri*, while sentences of conversation prefer linguistic elements that are associated with communicative function, such as the verb *haberi* (to show respect to the addressee) and expressive function, such as the particle *kana* and the interjection *ana*. In contrast, sentences of narration disprefer lexical items of communicative function as well as expressive function, whereas sentences of conversation disprefer linguistic elements to indicate tense or aspect. Short poems, on the other hand, have their own linguistic preference and prefer lexical items that indicate nature the so-called poetic words, such as *hototogisu* (a cuckoo), *shika* (deer) and *matsu* (pine trees). The targeted lexical items (the auxiliary verbs for speculation, particles for exclamation, interjections) are all classified as LFq (narration) as well as HFq (conversation) due to their association with subjectivity. Based on the findings from the LLR analysis, I then performed concordance research on the text of *The Tale of Genji* to identify the targeted lexical items rendered in sentences of narration. Although lexical items of HFq (conversation), as the LLR analysis demonstrates, are not likely to appear in sentences of narration, the concordance plot research shows that there are cases when a targeted lexical item appears in narration of *kaimami* scenes of the tale.

In Chapter 4, I examined sentences of direct representation of a character's consciousness appearing in *kaimami* scenes from five different chapters of *The Tale of Genji* (Chapter 3 *Utsusemi*, Chapter 5 *Wakamurasaki*, Chapter 28 *Nowaki*, Chapter 34 *Wakana: jo*, and Chapter 46 *Shiigamoto*). This qualitative analysis has identified three major effects of sentences of represented internal states in the text. Thoughts and emotions of a character, which may be conveyed through the use of a targeted lexical item, are often cited in the text (Matsushita, 2006). That is, a character's perceptions and emotions are often followed by a citational marker, such as the particles *to* and *tote*. However, a citational marker is omitted in

sentences of represented internal states (Mitani, 1994). In other words, without any interference such as a citational marker, sentences of represented internal states render a character's consciousness or subjectivity directly in sentences of narration, which lends vividness to the sentence. This vividness added in sentences of represented internal states, therefore, contributes to making the sentence read as if the reader were experiencing the scene or the feeling of a character on his own. In addition, sentences of represented internal states limit the visual or auditory information given to the reader by presenting information only available to a character who is experiencing it. That is, SRIS limit the information provided in the visual or auditory description in the text. This limited information of perception plays an important role. Firstly, it highlights a piece of information provided in the text, which may play a crucial role in the story. Secondly, providing a limited amount of visual or auditory information indicates the hidden information of a scene. In other words, the unseen or unheard may also be indicated in the text. Indicating the hidden information makes it possible to add depth to the description of a scene. SRIS, thus, create various effects by rendering the character's consciousness, perceptions, and emotions directly in the text.

5.2. Implication of the Study

Previous studies on sentences of represented internal states in classical Japanese as well as in modern Japanese (e.g., Ikegami, 1986; Itoi, 1992; Takahashi, 1992) claim the uniqueness of the Japanese language based on the high frequency of direct representation of a character's consciousness rendered in Japanese texts, and often state that this is possible due to the unique characteristics of the Japanese language. However, sentences of represented internal states are, in fact, found not only in *The Tale of Genji* or Japanese literature but also in literary works of other languages. Due to its functional and linguistic uniqueness, sentences of direct representation of a character's consciousness have also attracted the attention of literary scholars as well as linguists outside of Japanese literature/ language.

Though Prince (1987/ 2003) says it is difficult to determine linguistic elements to create this type of sentence, Banfield (1982), by exploring various types of sentences of represented internal states in literary texts of English with reference to German and French, finds that linguistic elements associated with subjectivity contribute to creating sentences of represented internal states. That is, there is a common linguistic quality to create this phenomenon in texts across different languages: subjectivity. As the current dissertation has demonstrated, sentences of represented internal states in the text of *The Tale of Genji* are also rendered by using linguistic elements associated with subjectivity. Across various languages, realization of certain features such as subjectivity may vary, and certain effects may be achieved by manipulating grammar in different ways. Therefore, it is crucial to investigate not only equivalent lexical items but also common linguistic functions rendered in the text.

5.3. Suggestions for Future Research

Systematic investigation from a linguistic point of view on other issues in the textual analysis of *The Tale of Genji* would expand our knowledge of the language in literary texts as well as provide a new perspective to existing issues in the textual analysis of the tale. For instance, those sentences of represented internal states that are examined in this dissertation are only one type of linguistic structure that has been identified in previous textual analyses of the tale. That is, there is another prominent structure proposed in the previous studies when a character's consciousness is rendered with citation, yet necessary honorific expressions are omitted. Combining the quantitative and qualitative analyses may provide a new perspective as well as statistical support to the claim of this structure. Also, it would be interesting to investigate *kaimami* scenes in other literary works written in the Heian period and analyze if sentences of represented internal states appear in the text and provide the same effects found in *The Tale of Genji*.

This quantitative method may also linguistically contribute to providing a new

perspective on a major discussion in the textual analysis of *The Tale of Genji*: the existence of the narrator. Previous studies on textual analysis claim that the use of honorifics indicates the existence of the narrator in the text of the tale and reveals her social status. The LLR analysis conducted in Chapter 3, in fact, shows that linguistic elements of honorifics show a distinctive distribution in the text type of narration in *The Tale of Genji* as opposed to the other 15 literary works from the Heian period. Honorific expressions of respect forms, such as *tamau* (to show respect) and *owasu* (to be), are classified as highest frequency words for narration as well as lowest frequency words for conversation in the text of the tale, whereas these honorific elements are classified as lowest frequency words for narration as well as highest frequency words for conversation in the other 15 literary works. This distributional difference of honorific expressions between the tale and the other literary works may suggest that the text of *The Tale of Genji* is rendered in a way that it feels like there exists the narrator in the text.

APPENDIX A

LIST OF HIGHEST AND LOWEST FREQUENCY WORDS

List of Highest Frequency Words for Narration in 16 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	nado (など)	particle (adverbial)	5666	1148	1	741.214
2	te (て)	particle (conjective)	21045	7416	1167	688.034
3	tari (たり: perfective)	auxiliary verb	7376	2114	93	535.669
4	to (と)	particle (case)	14605	4819	1322	421.919
5	iu (言ふ)	verb	4543	1121	194	350.321
6	ni (に)	particle (conjective)	2983	719	101	252.027
7	okashi (をかし)	adjective	1310	178	5	242.937
8	tamau (給ふ: respect)	verb	13986	6606	3	237.855
9	uchi (うち)	prefix	1801	310	67	225.761
10	obosu (思す)	verb	3021	944	0	199.430
11	tote (とて)	particle (case)	2192	541	69	181.041
12	ba (ば)	particle (conjective)	8471	3045	835	167.333
13	ge (気)	suffix	1722	437	11	159.276
14	ri (り)	auxiliary verb	3915	1274	232	151.405
15	notamau (宣ふ)	verb	1569	399	0	150.432
16	domo (共)	suffix	1828	514	2	147.093
17	otoko (男)	noun	749	115	0	128.559
18	iru (居る)	verb	915	136	40	126.306
19	kaku (書く)	verb	804	133	11	119.619
20	iru (入る)	verb	1246	278	40	118.742
21	izu (出ず)	verb	2422	744	154	104.443
22	ito (いと)	adverb	4954	2270	13	94.401
23	fusu (伏す)	verb	435	45	9	87.825
24	hitobito (人々)	noun	961	256	2	84.312
25	naku (泣く)	verb	532	43	48	82.789
26	keri (けり)	auxiliary verb	6194	2420	627	82.498
27	rei (例)	noun	706	158	0	81.214
28	yaru (遣る)	verb	793	144	56	76.933
29	warau (笑ふ)	verb	421	61	0	75.781
30	tsutsu (つつ)	particle (conjective)	1022	136	166	75.373
31	onna (女)	noun	907	255	0	73.419
32	kehai (気配)	noun	374	51	0	70.571
33	hiku (引く)	verb	767	162	43	67.780
34	sama (様)	noun	1670	640	7	62.921
35	do (ど)	particle (conjective)	2736	1083	176	52.321
36	tsuzuku (続く)	verb	315	50	1	51.580
37	sashi (差し)	prefix	528	122	11	51.303
38	kikoyu (聞こゆ)	verb	3162	1509	22	46.067
39	garu (がる)	suffix	298	52	0	45.474
40	shiroshi (白し)	adjective	222	24	3	45.438
41	utsukushi (美し)	adjective	238	32	0	45.379
42	kichoo (几帳)	noun	200	23	0	42.460
43	kaeshi (返し)	noun	210	25	3	40.287
44	kimi (君)	noun	1026	350	43	40.028
45	hodo (程)	noun	2125	909	75	39.344
46	miya (宮)	noun	1003	383	2	38.873
47	keshiki (気色)	noun	829	281	21	37.502
48	ka (日)	suffix	329	70	7	35.598
49	nonoshiru (罵る)	verb	184	25	0	34.807
50	gimi (君)	suffix	738	266	0	34.383

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	omoshiroshi (面白し)	adjective	238	44	0	34.250
52	fumi (文)	noun	545	161	14	33.785
53	meku (めく)	suffix	258	50	2	33.759
54	monogatari (物語)	noun	309	72	0	33.640
55	kata (方)	noun	1352	509	79	33.357
56	aware (哀れ)	noun	1133	402	75	31.858
57	katarau (語らう)	verb	278	54	9	31.433

List of Lowest Frequency Words for Narration in 16 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	haberi (侍り)	verb	216	3692	2	-2329.708
2	koso (こそ)	particle (bound)	1083	2440	340	-711.341
3	mu (む)	auxiliary verb	3563	4544	762	-628.743
4	namu (なむ)	particle (bound)	898	2301	2	-588.267
5	ki (き)	auxiliary verb	2503	3365	665	-552.037
6	tamau (給ふ: humble)	verb	11	642	0	-469.395
7	ya (や)	particle (bound)	1643	2082	776	-445.245
8	wa (は)	particle (bound)	9392	7823	1934	-426.295
9	ka (か)	particle (bound)	825	1307	386	-334.296
10	kimi (君)	pronoun	5	66	266	-244.837
11	kana (かな)	particle (final)	333	582	294	-228.976
12	na (な)	particle (final)	66	261	89	-150.695
13	na (な)	adverb	7	186	32	-150.144
14	owashimasu (おわします)	verb	481	884	0	-149.408
15	ji (じ)	auxiliary verb	349	541	187	-146.563
16	so (そ)	particle (final)	8	181	32	-143.792
17	tomo (とも)	particle (conjective)	169	362	131	-141.865
18	tsu (つ)	auxiliary verb	1401	1515	232	-136.041
19	moosu (申す)	verb	523	894	2	-135.661
20	nari (なり: hearsay)	auxiliary verb	247	531	51	-135.474
21	ima (今)	noun	658	865	144	-125.520
22	sasu (さす)	auxiliary verb	931	1259	1	-120.391
23	ramu (らむ)	auxiliary verb	527	498	355	-117.209
24	koto (事)	noun	4161	3723	207	-109.890
25	mi (身)	noun	388	384	294	-105.951
26	beshi (べし)	auxiliary verb	2472	2278	245	-100.779
27	nado (など)	adverb	107	293	33	-97.388
28	daijin (大臣)	noun	36	213	0	-96.677
29	yo (世)	noun	1128	1056	303	-96.058
30	zo (ぞ)	particle (bound)	2246	1527	778	-94.147
31	ana (あな)	interjection	65	243	15	-93.931
32	su (す)	auxiliary verb	1697	1810	24	-91.357
33	toshi (年)	noun	16	161	0	-89.012
34	uketamawaru (承る)	verb	23	176	0	-88.856
35	makaru (罷る)	verb	29	186	0	-87.508
36	ga (が)	particle (case)	1191	867	504	-84.463
37	dono (殿)	suffix	256	481	0	-83.926
38	nani (何)	pronoun	312	436	106	-83.882
39	yado (宿)	noun	7	8	121	-82.168
40	maro (麻呂)	pronoun	1	98	3	-76.229
41	meri (めり)	auxiliary verb	761	911	28	-70.925
42	ide (いで)	interjection	33	156	4	-65.691
43	koko (此処)	pronoun	164	313	15	-62.440
44	ikani (如何に)	adverb	366	481	59	-62.187
45	haya (早)	adverb	3	84	6	-61.696
46	ta (誰)	pronoun	10	87	22	-61.655
47	o (を)	particle (conjective)	114	174	92	-61.210
48	yo (よ)	particle (final)	148	279	21	-58.140
49	kaku (斯く)	adverb	927	975	50	-55.251
50	aki (秋)	noun	111	37	212	-54.759

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	sa (然)	adverb	571	673	35	-54.168
52	mi (み)	suffix	3	4	75	-53.243
53	ichi (一)	noun	77	202	0	-52.597
54	hana (花)	noun	265	90	322	-52.550
55	haru (春)	noun	83	53	156	-52.210
56	iza (いざ)	interjection	2	51	23	-51.952
57	kyoo (今日)	noun	190	195	137	-51.845
58	mashi (まし)	auxiliary verb	276	254	167	-51.781
59	ikaga (如何)	adverb	209	324	27	-51.358
60	sari (然り)	verb	930	1001	7	-50.643
61	daijoo (太政)	noun	5	81	0	-50.333
62	tanomu (頼む)	verb	87	106	104	-50.162
63	kashi (かし)	particle (final)	513	626	14	-49.781
64	ikade (如何で)	adverb	275	383	30	-49.273
65	shiru (知る)	verb	1100	796	337	-46.837
66	namu (なむ)	particle (final)	26	26	91	-46.043
67	kore (此れ)	pronoun	448	543	22	-45.231
68	kampaku (関白)	noun	5	73	0	-44.380
69	sode (袖)	noun	93	24	181	-44.151
70	kashikoshi (賢し)	adjective	80	184	5	-44.094
71	yama (山)	noun	208	105	225	-43.695
72	nanigashi (某)	pronoun	8	79	0	-43.428
73	tenno (天皇)	noun	4	69	0	-43.408
74	inochi (命)	noun	53	109	43	-43.055
75	dani (だに)	particle (adverbial)	474	442	138	-42.769
76	au (会う)	verb	176	105	188	-42.234
77	toshi (年)	noun	235	248	101	-40.712
78	tadaima (只今)	noun	67	165	0	-40.266
79	tsuyu (露)	noun	56	24	121	-37.263
80	choo (てふ)	verb	2	1	52	-35.750
81	guu (宮)	suffix	2	53	0	-35.750
82	monosu (ものす)	verb	391	480	0	-35.711
83	onore (己)	pronoun	10	69	3	-35.513
84	shika (然)	adverb	25	85	12	-34.777
85	nu (ぬ)	auxiliary verb	3753	2357	736	-34.749
86	oboshimesu (思し召す)	verb	93	184	0	-34.509
87	oba (をば)	particle (case)	264	282	68	-31.779
88	ware (我)	pronoun	428	282	216	-31.664
89	domo (ども)	particle (conjective)	139	98	129	-31.647
90	kou (恋う)	verb	43	12	105	-31.560
91	ushi (憂し)	adjective	143	71	159	-31.211

List of Highest Frequency Words for Conversation in 16 literary works

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
1	haberi (侍り)	verb	216	3692	2	2938.189
2	namu (なむ)	particle (bound)	898	2301	2	893.331
3	koso (こそ)	particle (bound)	1083	2440	340	651.451
4	tamau (給ふ: humble)	verb	11	642	0	578.251
5	mu (む)	auxiliary verb	3563	4544	762	565.462
6	ki (き)	auxiliary verb	2503	3365	665	427.386
7	koto (事)	noun	4161	3723	207	268.959
8	owashimasu (おわします)	verb	481	884	0	254.007
9	wa (は)	particle (bound)	9392	7823	1934	245.111
10	sasu (さす)	auxiliary verb	931	1259	1	244.796
11	moosu (申す)	verb	523	894	2	236.043
12	su (す)	auxiliary verb	1697	1810	24	224.204
13	ka (か)	particle (bound)	825	1307	386	170.776
14	beshi (べし)	auxiliary verb	2472	2278	245	155.556
15	ya (や)	particle (bound)	1643	2082	776	154.326
16	nari (なり: hearsay)	auxiliary verb	247	531	51	147.451
17	dono (殿)	suffix	256	481	0	141.348
18	tsu (つ)	auxiliary verb	1401	1515	232	138.914
19	meri (めり)	auxiliary verb	761	911	28	135.829
20	otodo (大臣)	noun	36	213	0	129.619
21	sari (然り)	verb	930	1001	7	127.938
22	uketamawaru (承る)	verb	23	176	0	116.731
23	makaru (罷る)	verb	29	186	0	116.510
24	toshi (年)	noun	16	161	0	115.001
25	ima (今)	noun	658	865	144	112.745
26	kaku (斯く)	adverb	927	975	50	107.113
27	ana (あな)	interjection	65	243	15	106.192
28	na (な)	adverb	7	186	32	103.705
29	so (そ)	particle (final)	8	181	32	98.542
30	kashi (かし)	particle (final)	513	626	14	98.264
31	ko (是)	pronoun	1940	1622	77	95.994
32	nado (など)	adverb	107	293	33	95.706
33	sa (然)	adverb	571	673	35	93.036
34	koko (此処)	pronoun	164	313	15	85.075
35	ide (いで)	interjection	33	156	4	82.113
36	maro (麻呂)	pronoun	1	98	3	81.757
37	kore (此れ)	pronoun	448	543	22	80.981
38	monosu (ものす)	verb	391	480	0	79.817
39	ichi (一)	noun	77	202	0	79.667
40	yo (よ)	particle (final)	148	279	21	70.937
41	ikani (如何に)	adverb	366	481	59	68.969
42	na (な)	particle (final)	66	261	89	67.931
43	ikade (如何で)	adverb	275	383	30	66.165
44	ikaga (如何)	adverb	209	324	27	64.413
45	saburau (侍ふ)	verb	408	456	0	63.801
46	daijoo (太政)	noun	5	81	0	63.703
47	tadaima (只今)	noun	67	165	0	61.976
48	kashikoshi (賢し)	adjective	80	184	5	61.975
49	ji (じ)	auxiliary verb	349	541	187	60.953
50	haya (早)	adverb	3	84	6	58.877

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
51	tomo (とも)	article (conjective	169	362	131	58.401
52	oboshimesu (思し召す)	verb	93	184	0	56.940
53	kanpaku (関白)	noun	5	73	0	56.382
54	nanigashi (某)	pronoun	8	79	0	56.164
55	tenno (天皇)	noun	4	69	0	54.822
56	tatematsuru (奉る)	verb	1510	1146	4	53.797
57	kana (かな)	particle (final)	333	582	294	53.309
58	nani (何)	pronoun	312	436	106	52.867
59	usu (失す)	verb	116	199	1	52.396
60	sore (其れ)	pronoun	294	352	20	49.529
61	guu (宮)	suffix	2	53	0	44.622
62	yo (世)	noun	1128	1056	303	44.513
63	nyuudoo (入道)	noun	59	129	0	43.836
64	mikado (帝)	noun	210	259	0	43.390
65	kurai (位)	noun	42	110	0	43.322
66	so (其)	pronoun	755	649	29	42.312
67	maji (まじ)	auxiliary verb	388	387	1	42.143
68	yoshi (良し)	adjective	346	361	8	41.622
69	onore (己)	pronoun	10	69	3	40.248
70	tooguu (東宮)	noun	8	59	0	38.686
71	zoo (贈)	noun	0	38	0	36.849
72	nanigoto (何事)	noun	182	224	3	36.344
73	koo (斯う)	adverb	228	256	1	35.817
74	ta (誰)	pronoun	10	87	22	35.218
75	suke (出家)	noun	0	36	0	34.909
76	sarani (更に)	adverb	224	259	14	34.334

List of Lowest Frequency Words for Conversation in 16 literary works

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
1	to (と)	particle (case)	14605	4819	1322	-403.042
2	nado (など)	particle (adverbial)	5666	1148	1	-384.633
3	te (て)	particle (conjective)	21045	7416	1167	-377.315
4	iu (言ふ)	verb	4543	1121	194	-236.719
5	tari (たり: perfect)	auxiliary verb	7376	2114	93	-233.850
6	uchi (打ち)	prefix	1801	310	67	-170.569
7	ba (ば)	particle (conjective)	8471	3045	835	-170.036
8	tsutsu (つつ)	particle (conjective)	1022	136	166	-160.763
9	ni (に)	particle (conjective)	2983	719	101	-158.897
10	okashi (をかし)	adjective	1310	178	5	-152.123
11	tote (とて)	particle (case)	2192	541	69	-109.857
12	naku (泣く)	verb	532	43	48	-105.732
13	iru (居る)	verb	915	136	40	-104.653
14	ri (り)	auxiliary verb	3915	1274	232	-97.459
15	keri (けり)	auxiliary verb	6194	2420	627	-82.395
16	iru (入る)	verb	1246	278	40	-77.101
17	kaku (書く)	verb	804	133	11	-75.970
18	otoko (男)	nou	749	115	0	-75.302
19	yaru (遣る)	verb	793	144	56	-75.204
20	izu (出ず)	verb	2422	744	154	-75.065
21	ge (気)	suffix	1722	437	11	-74.286
22	obosu (思す)	verb	3021	944	0	-66.364
23	namida (涙)	noun	206	19	118	-66.300
24	fusu (伏す)	verb	435	45	9	-66.178
25	notamau (宣ふ)	verb	1569	399	0	-65.832
26	hana (花)	noun	265	90	322	-59.103
27	naku (鳴く)	verb	101	12	160	-58.826
28	domo (共)	suffix	1828	514	2	-58.208
29	hiku (引く)	verb	767	162	43	-55.853
30	nagamu (眺む)	verb	260	31	57	-48.503
31	warau (笑ふ)	verb	421	61	0	-45.236
32	sode (袖)	noun	93	24	181	-45.203
33	aki (秋)	noun	111	37	212	-43.644
34	miyu (見ゆ)	verb	1254	400	156	-43.451
35	kehai (気配)	noun	374	51	0	-42.891
36	koe (声)	noun	359	79	102	-40.247
37	ei (例)	noun	706	158	0	-39.505
38	chiru (散る)	verb	92	18	122	-36.221
39	hitobito (人々)	noun	961	256	2	-35.913
40	iku (行く)	verb	530	197	263	-35.218
41	sawagu (騒ぐ)	verb	266	38	25	-34.864
42	fuku (吹く)	verb	218	51	116	-33.784

List of Highest Frequency Words for Short Poem in 16 literary works

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
1	kimi (君)	pronoun	5	66	266	464.554
2	hana (花)	noun	265	90	322	355.205
3	aki (秋)	noun	111	37	212	284.362
4	sode (袖)	noun	93	24	181	249.557
5	yado (宿)	noun	7	8	121	237.647
6	ya (や)	particle (bound)	1643	2082	776	223.988
7	zo (ぞ)	particle (bound)	2246	1527	778	220.734
8	yama (山)	noun	208	105	225	218.929
9	naku (鳴く)	verb	101	12	160	213.645
10	no (の)	particle (case)	25599	13996	4333	206.292
11	ramu (らむ)	auxiliary verb	527	498	355	197.347
12	wa (わ)	pronoun	533	224	303	193.206
13	haru (春)	noun	83	53	156	191.558
14	ga (が)	particle (case)	1191	867	504	185.864
15	mi (身)	noun	388	384	294	179.123
16	au (会う)	verb	176	105	188	175.230
17	tsuyu (露)	noun	56	24	121	165.463
18	ushi (憂し)	adjective	143	71	159	157.617
19	kou (恋ふ)	verb	43	12	105	155.058
20	iku (行く)	verb	530	197	263	152.552
21	mi (み)	suffix	3	4	75	151.773
22	kana (かな)	particle (final)	333	582	294	151.064
23	chiru (散る)	verb	92	18	122	147.630
24	shi (し)	particle (adverbial)	106	34	124	135.030
25	namu (なむ)	particle (final)	26	26	91	130.742
26	kaze (風)	noun	189	65	143	119.121
27	kiyu (消ゆ)	verb	69	24	98	115.919
28	iro (色)	noun	230	67	150	115.576
29	choo (ちょう)	verb	2	1	52	109.394
30	nami (波)	noun	46	15	82	107.328
31	zu (ず)	auxiliary verb	6502	3697	1277	105.773
32	domo (ども)	particle (conjective)	139	98	129	104.944
33	ware (我)	pronoun	428	282	216	104.419
34	ama (海女)	noun	16	4	61	101.928
35	wa (は)	particle (bound)	9392	7823	1934	101.749
36	koi (恋)	noun	8	1	52	97.722
37	hototogisu (時鳥)	noun	30	8	67	96.484
38	koishi (恋し)	adjective	127	45	106	93.782
39	namida (涙)	noun	206	19	118	93.463
40	fu (経)	verb	198	107	134	92.490
41	kawa (川)	noun	75	18	84	92.356
42	nagaru (流る)	verb	32	6	64	90.792
43	matsu (待つ)	verb	188	92	126	88.698
44	kage (影)	noun	59	23	78	87.959
45	ku (来)	verb	914	412	286	86.999
46	ka (か)	particle (bound)	825	1307	386	85.878
47	sakurabana (桜花)	noun	0	0	37	84.893
48	saku (咲く)	verb	68	21	78	84.465
49	tanomu (頼む)	verb	87	106	104	83.991
50	mashi (まし)	auxiliary verb	276	254	167	83.923

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
51	tayu (絶ゆ)	verb	214	95	124	79.152
52	omooyu (思ほゆ)	verb	21	6	53	78.914
53	fuku (吹く)	verb	218	51	116	78.815
54	kyoo (今日)	noun	190	195	137	78.121
55	kumo (雲)	noun	35	11	60	77.612
56	furu (降る)	verb	213	56	114	76.319
57	na (名)	noun	126	76	100	75.874
58	uguisu (鶯)	noun	25	4	51	73.377
59	koromo (衣)	noun	10	7	45	72.709
60	mizu (水)	noun	120	34	87	72.620
61	shiru (知る)	verb	1100	796	337	72.193
62	ura (浦)	noun	25	12	53	71.040
63	se (瀬)	noun	5	7	41	70.217
64	wakaru (別る)	verb	48	24	63	68.161
65	miru (見る)	verb	3147	1446	623	67.415
66	matsu (松)	noun	57	14	62	67.010
67	nuru (濡る)	verb	46	12	58	66.966
68	kumoi (雲居)	noun	7	3	38	66.494
69	rashi (らし)	auxiliary verb	0	1	30	65.058
70	momijiba (紅葉)	noun	1	1	30	62.461
71	tamoto (袂)	noun	6	2	34	60.760
72	karakoromo (唐衣)	noun	0	1	28	60.528
73	sora (空)	noun	216	56	100	58.848
74	ama (天)	noun	6	2	33	58.651
75	yuki (雪)	noun	141	44	83	58.268
76	no (野)	noun	27	12	46	57.124
77	nobe (野辺)	noun	8	9	37	56.768
78	ma (間)	noun	115	41	75	55.632
79	tsuki (月)	noun	308	115	122	55.391
80	shirakumo (白雲)	noun	0	0	24	55.064
81	ne (音)	noun	120	70	81	54.490
82	ji (じ)	auxiliary verb	349	541	187	54.451
83	beranari (べらなり)	auxiliary verb	0	1	25	53.741
84	akikaze (秋風)	noun	4	2	29	52.993
85	ashihiki (足引き)	noun	0	0	23	52.769
85	shiratsuyu (白露)	noun	0	0	23	52.769
87	tazunu (訪ぬ)	verb	31	23	48	52.344
88	nu (ぬ)	auxiliary verb	3753	2357	736	52.052
89	o (を)	particle (case)	9599	5034	1521	51.847
90	omoi (思ひ)	noun	100	63	73	51.170
91	ominaeshi (女郎花)	noun	15	1	33	49.856
92	yamabe (山辺)	noun	3	1	26	49.677
93	kusa (草)	noun	44	17	48	49.100
94	moyu (燃ゆ)	verb	7	2	29	49.046
95	tomo (とも)	particle (conjective)	169	362	131	48.651
96	shigure (時雨)	noun	23	7	38	48.583
97	shimo (霜)	noun	17	3	34	48.376
98	o (を)	particle (conjective)	114	174	92	46.903
99	oosaka (あふさか)	noun	12	3	31	46.862
100	oru (折る)	verb	98	33	62	45.425

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
101	nu (寝)	verb	179	48	80	45.164
102	keburi (煙)	noun	28	14	40	45.125
103	harugasumi (春霞)	noun	0	1	21	44.711
104	waku (分く)	verb	178	112	90	44.448
105	ni (に)	particle (case)	16898	9033	2468	44.346
106	wasuru (忘る)	verb	166	119	89	44.259
107	ka (香)	noun	53	13	46	43.910
108	wabu (詫ぶ)	verb	115	37	65	43.837
109	tou (訪ふ)	verb	23	10	36	43.231
110	ou (生ふ)	verb	60	51	56	43.088
111	ume (梅)	noun	49	9	43	42.547
112	yume (夢)	noun	124	88	75	42.501
113	koyu (越ゆ)	verb	26	15	38	42.306
114	kami (神)	noun	61	54	56	41.946
115	oki (沖)	noun	1	0	19	40.207
116	shiranami (白波)	noun	1	1	20	40.193
117	goto (毎)	suffix	82	26	53	39.935
118	ato (跡)	noun	37	20	41	39.868
119	moga (もが)	particle (final)	31	20	39	39.267
120	namidagawa (涙川)	noun	0	0	17	39.003
121	ha (葉)	noun	52	12	42	38.653
122	yoshino (吉野)	noun	5	1	22	38.223
123	karu (刈る)	verb	15	4	28	37.918
124	ki (木)	noun	43	8	38	37.705
125	na (な)	particle (final)	66	261	89	37.628
126	furusato (古里)	noun	21	5	30	36.909
127	yoso (余所)	noun	40	38	44	36.681
128	amanogawa (天の川)	noun	3	3	21	36.143
129	ne (根)	noun	13	6	27	36.079
130	ki (き)	auxiliary verb	2503	3365	665	35.879
131	chihayaburu (千早振る)	noun	1	0	17	35.712
132	kesa (今朝)	noun	23	26	36	35.435
133	musubu (結ぶ)	verb	44	9	37	35.351
134	kotonoha (言の葉)	noun	44	13	38	35.312
135	mine (峰)	noun	19	10	30	35.178
136	yo (世)	noun	1128	1056	303	35.035
137	koe (声)	noun	359	79	102	34.865
138	shirayuki (白雪)	noun	0	0	15	34.414
139	utsurou (移ろふ)	verb	43	25	40	34.308
140	kanashi (悲し)	adjective	348	149	109	33.996
141	wakare (別れ)	noun	23	27	35	33.494
142	toko (床)	noun	0	1	16	33.469
143	chiyo (千代)	noun	3	2	19	33.245
143	ho (穂)	noun	4	1	19	33.245
145	niou (匂ふ)	verb	62	7	39	32.556
146	tatsuta (竜田)	noun	5	1	19	32.010
147	hanasusuki (花薄)	noun	2	1	17	31.852
148	seki (関)	noun	31	2	29	31.439
149	kari (雁)	noun	15	3	24	31.306
150	momiji (紅葉)	noun	57	16	39	31.291

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
151	hisakata (久方)	noun	1	0	15	31.229
152	karu (枯る)	verb	29	5	29	30.961
153	arashi (嵐)	noun	7	0	19	30.886
154	tare (誰)	pronoun	221	179	92	30.868
155	chidori (千鳥)	noun	4	0	17	30.378
156	yo (夜)	noun	518	175	129	30.230
157	ayanashi (文無し)	adjective	2	10	21	30.162
158	tsumu (摘む)	verb	12	11	25	29.972
159	watastsuumi (わたつうみ)	noun	0	0	13	29.825
160	nomi (のみ)	particle (adverbial)	771	351	178	29.556
161	okuyama (奥山)	noun	1	1	15	29.194
161	hitsu (漬つ)	verb	1	1	15	29.194
163	toshi (年)	noun	235	248	101	29.153
164	e (方)	noun	18	13	27	29.138
165	nageki (嘆き)	noun	16	26	30	29.051
166	ayamegusa (菖蒲草)	noun	1	0	14	28.993
166	ne (嶺)	noun	1	0	14	28.993
168	miyako (都)	noun	28	35	35	28.869
169	michi (道)	noun	178	113	74	28.633
170	shika (鹿)	noun	7	5	20	28.268
171	kawaku (乾く)	verb	3	3	17	27.921
172	kokoro (心)	noun	2088	1178	394	27.912
173	tama (魂)	noun	2	1	15	27.562
174	koromode (衣手)	noun	0	0	12	27.531
174	moru (守る)	verb	0	0	12	27.531
176	suminoe (住の江)	noun	1	1	14	27.012
176	yorozuyo (万代)	noun	2	0	14	27.012
176	yawa (夜半)	noun	2	0	14	27.012
179	chigiru (契る)	verb	56	18	36	26.942
180	tama (玉)	noun	40	36	36	26.389
181	uu (植う)	verb	34	11	29	26.387
182	hama (浜)	noun	27	3	25	26.368
183	sato (里)	noun	83	30	43	26.133
184	koma (駒)	noun	12	13	23	25.511
185	manima (随)	noun	3	0	14	25.431
186	sasagani (細蟹)	noun	0	0	11	25.236
186	tagitsu (滾つ)	verb	0	0	11	25.236
186	michinoku (ミチノク)	noun	0	0	11	25.236
189	kasumi (霞)	noun	19	7	23	25.017
190	shio (潮)	noun	3	4	16	24.881
191	aru (荒る)	verb	58	15	34	24.593
192	yamazato (山里)	noun	44	26	33	24.099
193	iku (幾)	noun	9	19	23	24.075
194	hiru (干る)	verb	6	2	16	23.953
195	fuchi (淵)	noun	4	2	15	23.886
196	kiku (菊)	noun	38	11	28	23.548
197	shita (下)	noun	103	36	45	23.266
198	kishi (岸)	noun	17	3	20	23.087
199	tomu (止む)	verb	30	8	25	23.041
200	azusayumi (梓弓)	noun	0	0	10	22.942

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
200	amagumo (天雲)	noun	0	0	10	22.942
200	sohotsu (そほつ)	verb	0	0	10	22.942
200	watatsumi (海神)	noun	0	0	10	22.942
204	oozora (大空)	noun	4	1	14	22.929
205	kaku (掛く)	verb	349	136	92	22.340
206	akigiri (秋霧)	noun	1	0	11	22.311
206	shiratama (白玉)	noun	1	0	11	22.311
206	miyama (深山)	noun	1	0	11	22.311
209	aku (飽く)	verb	158	45	54	22.168
210	kazu (数)	noun	94	73	48	21.691
211	shigeshi (繁し)	adjective	83	24	38	21.611

List of Lowest Frequency Words for Short Poem in 16 literary works

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
1	tamau (給ふ: respect)	verb	13986	6606	3	-1320.724
2	oon (御)	prefix	4691	2759	2	-469.223
3	nado (など)	particle (adverbial)	5666	1148	1	-433.216
4	ito (いと)	adverb	4954	2270	13	-413.086
5	tari (たり: perfective)	auxiliary verb	7376	2114	93	-370.761
6	obosu (思す)	verb	3021	944	0	-255.270
7	haberi (侍り)	verb	216	3692	2	-241.270
8	te (て)	particle (conjective)	21045	7416	1167	-240.478
9	kikoyu (聞こゆ)	verb	3162	1509	22	-229.681
10	nari (なり: declarative)	auxiliary verb	15996	8919	989	-227.353
11	namu (なむ)	particle (bound)	898	2301	2	-195.885
12	su (す)	auxiliary verb	1697	1810	24	-155.872
13	tatematsuru (奉る)	verb	1510	1146	4	-153.969
14	koto (事)	noun	4161	3723	207	-145.720
15	domo (共)	suffix	1828	514	2	-141.200
16	sasu (さす)	auxiliary verb	931	1259	1	-135.582
17	mi (御)	prefix	2001	1141	26	-130.797
18	notamau (宣ふ)	verb	1569	399	0	-126.545
19	mairu (参る)	verb	1207	811	1	-124.578
20	sama (様)	noun	1670	640	7	-123.262
21	imiji (いみじ)	adjective	1148	603	1	-107.514
22	ge (気)	suffix	1722	437	11	-104.034
23	yoo (様)	noun	959	628	0	-102.022
24	sari (然り)	verb	930	1001	7	-99.948
25	owasu (おわす)	verb	839	556	0	-89.669
26	owashimasu (おわします)	verb	481	884	0	-87.739
27	moosu (申す)	verb	523	894	2	-82.530
28	ko (是)	pronoun	1940	1622	77	-81.633
29	miya (宮)	noun	1003	383	2	-80.574
30	okashi (をかし)	adjective	1310	178	5	-78.029
31	oboyu (覚ゆ)	verb	989	474	5	-76.494
32	onna(女)	noun	907	255	0	-74.681
33	hitobito (人々)	noun	961	256	2	-69.929
34	iu (言ふ)	verb	4543	1121	194	-69.315
35	ni (に)	particle (conjective)	2983	719	101	-64.826
36	gimi (君)	suffix	738	266	0	-64.520
37	hodo (程)	noun	2125	909	75	-60.098
38	otodo (大臣)	noun	541	394	0	-60.084
39	tokoro (所)	noun	942	489	12	-59.196
40	monosu (ものす)	verb	391	480	0	-55.969
41	saburau (侍ふ)	verb	408	456	0	-55.519
42	rei (例)	noun	706	158	0	-55.519
43	otoko (男)	noun	749	115	0	-55.519
44	tote (とて)	particle (case)	2192	541	69	-52.775
45	arisama (有り様)	noun	553	300	1	-50.374
46	ka (彼)	pronoun	669	390	6	-49.629
47	meri (めり)	auxiliary verb	761	911	28	-47.702
48	dono (殿)	suffix	256	481	0	-47.354
49	sukosi (少し)	adverb	624	227	2	-47.015
50	tachi (達)	suffix	446	264	0	-45.619

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
51	maji (まじ)	auxiliary verb	388	387	1	-45.443
52	raru (らる)	auxiliary verb	733	594	17	-45.071
53	in (院)	noun	369	325	0	-44.590
54	e (え)	adverb	747	477	14	-44.175
55	tamau (給ふ: humble)	verb	11	642	0	-41.955
56	nagon (納言)	noun	419	226	0	-41.441
57	mote (以て)	prefix	536	247	3	-40.081
58	kashi (かし)	particle (final)	513	626	14	-39.579
59	tono (殿)	noun	349	314	1	-38.379
60	tsuku (付く)	verb	1113	486	36	-35.092
61	kaku (斯く)	adverb	927	975	50	-34.903
62	itooshi (いとおし)	adjective	361	172	0	-34.242
63	so (其)	pronoun	755	649	29	-33.566
64	ri (り)	auxiliary verb	3915	1274	232	-33.522
65	kaku (書く)	verb	804	133	11	-33.360
66	zu (ず)	verb	268	247	0	-33.086
67	ayashi (怪し)	adjective	508	414	11	-32.549
68	warau (笑ふ)	verb	421	61	0	-30.965
69	kokochi (心地)	noun	881	342	24	-30.642
70	goran (御覧)	noun	234	242	0	-30.579
71	chuujoo (中将)	noun	366	109	0	-30.515
72	mikado (帝)	noun	210	259	0	-30.130
73	uchi (打ち)	prefix	1801	310	67	-29.448
74	keshiki (気色)	noun	829	281	21	-28.711
75	miko (御子)	noun	247	190	0	-28.073
76	iru (入る)	verb	1246	278	40	-28.013
77	tukoomatsuru (仕る)	verb	221	212	0	-27.816
78	kuchioshi (口惜し)	adjective	262	170	0	-27.752
79	kehai (気配)	noun	374	51	0	-27.302
80	kuruma (車)	noun	377	154	2	-27.264
81	koo (斯う)	adverb	228	256	1	-27.148
82	nichi (日)	noun	250	166	0	-26.724
83	medetashi (めでたし)	adjective	371	150	2	-26.655
84	mesu (召す)	verb	306	105	0	-26.403
85	kitanokata (北の方)	noun	290	118	0	-26.210
86	o (御)	prefix	367	145	2	-26.107
87	ari (有り)	verb	5703	3278	491	-26.037
88	toru (取る)	verb	856	466	33	-25.845
89	yoshi (良し)	adjective	346	361	8	-25.645
90	taishoo (大将)	noun	286	105	0	-25.117
91	mina (皆)	noun	521	176	8	-25.100
92	geni (げに)	adverb	381	254	6	-25.028
93	tada (唯)	adverb	742	426	27	-24.912
94	yorozu (万)	noun	320	125	1	-24.715
95	monogatari (物語)	noun	309	72	0	-24.475
96	ooshi (多し)	adjective	567	221	12	-24.035
97	shooshoo (少将)	noun	296	78	0	-24.025
98	kokorogurushi (心苦し)	adjective	241	127	0	-23.640
99	yomu (読む)	verb	419	124	4	-23.511
100	sate (さて)	conjunction	212	212	1	-23.407

ranking	lexical item	parts of speech	narration	conversation	short poem	LLR value
101	mata (また)	conjunction	201	151	0	-22.612
102	hazukashi (恥ずかし)	adjective	321	89	1	-22.537
103	garu (がる)	suffix	298	52	0	-22.483
104	do (ど)	particle (conjective)	2736	1083	176	-22.366
105	kore (此れ)	pronoun	448	543	22	-22.056
106	juu (十)	noun	184	159	0	-22.033
107	chuu (中)	prefix	232	110	0	-21.969
108	notamawasu (宣はす)	verb	204	133	0	-21.648
109	yoku (良く)	adverb	191	140	0	-21.262
110	itoo (いたう)	adverb	259	72	0	-21.262
111	toshigoro (年頃)	noun	207	181	1	-21.171
112	sa (然)	adverb	571	673	35	-20.837
113	musume (娘)	noun	174	150	0	-20.813
114	kami (守)	noun	269	108	1	-20.489

List of Highest Frequency Words for Narration in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	tamau (給ふ: respect)	verb	11207	2691	0	563.586
2	to (と)	particle (case)	7816	1973	229	256.218
3	nado (など)	particle (adverbial)	3229	608	0	250.437
4	te (て)	particle (conjective)	11563	3414	283	244.551
5	ri (り)	auxiliary verb	2886	501	54	216.359
6	obosu (思す)	verb	2782	584	0	179.634
7	notamau (宣ふ)	verb	1065	139	0	129.135
8	tari (たり: perfective)	auxiliary verb	3379	909	31	113.698
9	oon (御)	prefix	3696	1086	0	104.675
10	uchi (打ち)	prefix	1169	191	19	96.822
11	ni (に)	particle (conjective)	1691	388	12	85.783
12	tsutsu (つつ)	particle (conjective)	707	79	23	77.192
13	okashi (をかし)	adjective	589	73	0	74.946
14	domo (共)	suffix	1129	235	0	73.723
15	kimi (君)	noun	731	115	10	64.824
16	ba (ば)	particle (conjective)	3966	1242	97	64.351
17	tote (とて)	particle (case)	939	176	15	63.827
18	iru (居る)	verb	411	30	10	63.650
19	izu (出ず)	verb	1484	358	20	63.232
20	naku (泣く)	verb	303	10	10	59.093
21	ge (気)	suffix	1118	259	1	58.760
22	hitobito (人々)	noun	604	107	0	50.973
23	fusu (伏す)	verb	209	8	0	50.340
24	gimi (君)	suffix	554	94	0	49.627
25	kehai (気配)	noun	331	35	0	48.174
26	tsuzuku (続く)	verb	225	16	0	42.244
27	utsukushi (美し)	adjective	178	8	0	40.697
28	miya (宮)	noun	782	183	0	40.518
29	iru (入る)	verb	613	117	10	40.314
30	kaku (書く)	verb	377	55	1	39.810
31	rei (例)	noun	449	79	0	38.248
32	owasu (おはす)	verb	645	145	0	36.333

List of Lowest Frequency Words for Narration in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	haberi (侍り)	verb	7	2408	0	-2299.695
2	namu (なむ)	particle (bound)	359	1416	0	-719.570
3	koso (こそ)	particle (bound)	561	1266	58	-479.155
4	tamau (給ふ: humble)	verb	3	498	0	-466.892
5	mu (む)	auxiliary verb	2329	2325	158	-314.544
6	ya (や)	particle (bound)	1046	1104	196	-219.236
7	ki (き)	auxiliary verb	1524	1395	195	-191.408
8	ka (か)	particle (bound)	457	616	71	-154.034
9	wa (は)	particle (bound)	5799	3644	343	-118.015
10	beshi (べし)	auxiliary verb	1719	1404	57	-102.328
11	kana (かな)	particle (final)	264	318	88	-93.704
12	nari (なり: hearsay)	auxiliary verb	115	262	0	-91.691
13	monosu (ものす)	verb	207	338	0	-84.112
14	yo (世)	noun	878	730	110	-82.096
15	na (な)	adverb	2	83	7	-79.320
16	so (そ)	particle (final)	2	82	7	-78.372
17	na (な)	particle (final)	34	126	18	-75.570
18	tomo (とも)	particle (conjective)	120	206	35	-75.134
19	koto (事)	noun	2754	1931	33	-69.354
20	ji (じ)	auxiliary verb	226	289	37	-68.959
21	tsu (つ)	auxiliary verb	795	697	41	-66.631
22	meri (めり)	auxiliary verb	472	504	6	-65.842
23	mi (身)	noun	274	272	76	-60.611
24	kimi (君)	pronoun	4	32	42	-59.208
25	makaru (罷る)	verb	4	65	0	-50.946
26	uketamawaru (承る)	verb	13	82	0	-50.821
27	ima (今)	noun	422	409	14	-46.235
28	nu (ぬ)	auxiliary verb	1894	1226	116	-45.746
29	ana (あな)	interjection	41	108	2	-43.821
30	ide (いで)	interjection	25	89	1	-43.347
31	nani (何)	pronoun	169	208	17	-42.247
32	kaku (斯く)	adverb	606	526	10	-42.207
33	nanigashi (某)	pronoun	4	54	0	-40.936
34	sutsu (捨つ)	verb	120	161	14	-37.565
35	ramu (らむ)	auxiliary verb	333	274	60	-36.545
36	hito (人)	noun	2339	1442	103	-35.708

List of Highest Frequency Words for Conversation in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	haberi (侍り)	verb	7	2408	0	2498.563
2	namu (なむ)	particle (bound)	359	1416	0	822.781
3	tamau (給ふ: humble)	verb	3	498	0	507.677
4	koso (こそ)	particle (bound)	561	1266	58	484.241
5	mu (む)	auxiliary verb	2329	2325	158	314.208
6	ki (き)	auxiliary verb	1524	1395	195	134.013
7	ya (や)	particle (bound)	1046	1104	196	132.679
8	beshi (べし)	auxiliary verb	1719	1404	57	125.536
9	ka (か)	particle (bound)	457	616	71	123.344
10	koto (事)	noun	2754	1931	33	113.339
11	nari (なり: hearsay)	auxiliary verb	115	262	0	108.878
12	wa (は)	particle (bound)	5799	3644	343	106.991
13	monosu (ものす)	verb	207	338	0	104.093
14	meri (めり)	auxiliary verb	472	504	6	85.232
15	tsu (つ)	auxiliary verb	795	697	41	71.586
16	na (な)	adverb	2	83	7	64.218
17	so (そ)	particle (final)	2	82	7	63.260
18	kaku (斯く)	adverb	606	526	10	57.300
19	uketamawaru (承る)	verb	13	82	0	57.058
20	makaru (罷る)	verb	4	65	0	56.126
21	ima (今)	noun	422	409	14	55.133
22	na (な)	particle (final)	34	126	18	54.786
23	yo (世)	noun	878	730	110	53.082
24	ji (じ)	auxiliary verb	226	289	37	52.441
25	sasu (さす)	auxiliary verb	541	464	0	51.077
26	tomo (とも)	particle (conjective)	120	206	35	49.660
27	ana (あな)	interjection	41	108	2	48.168
28	ide (いで)	interjection	25	89	1	47.959
29	sari (然り)	verb	611	493	0	46.406
30	nanigashi (某)	pronoun	4	54	0	45.213
31	maji (まじ)	auxiliary verb	287	286	1	43.227
32	nu (ぬ)	auxiliary verb	1894	1226	116	41.149
33	hito (人)	noun	2339	1442	103	40.642
34	kana (かな)	particle (final)	264	318	88	39.484
35	nani (何)	pronoun	169	208	17	38.845

List of Lowest Frequency Words for Conversation in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	tamau (給ふ: respect)	verb	11207	2691	0	-358.377
2	to (と)	particle (case)	7816	1973	229	-239.565
3	te (て)	particle (conjective)	11563	3414	283	-196.228
4	ri (り)	auxiliary verb	2886	501	54	-193.695
5	nado (など)	particle (adverbial)	3229	608	0	-178.718
6	obosu (思す)	verb	2782	584	0	-122.860
7	notamau (宣ふ)	verb	1065	139	0	-100.549
8	uchi (打ち)	prefix	1169	191	19	-85.370
9	tsutsu (つつ)	particle (conjective)	707	79	23	-83.489
10	tari (たり: perfective)	auxiliary verb	3379	909	31	-75.862
11	naku (泣く)	verb	303	10	10	-68.396
12	iru (居る)	verb	411	30	10	-65.656
13	ni (に)	particle (conjective)	1691	388	12	-61.435
14	okashi (をかし)	adjective	589	73	0	-58.835
15	kimi (君)	noun	731	115	10	-56.016
16	oon (御)	prefix	3696	1086	0	-54.677
17	tote (とて)	particle (case)	939	176	15	-54.551
18	domo (共)	suffix	1129	235	0	-50.611
19	ba (ば)	particle (conjective)	3966	1242	97	-49.177
20	izu (出ず)	verb	1484	358	20	-48.087
21	fusu (伏す)	verb	209	8	0	-43.160
22	ge (気)	suffix	1118	259	1	-38.645
23	kehai (気配)	noun	331	35	0	-38.631
24	hitobito (人々)	noun	604	107	0	-37.088
25	gimi (君)	suffix	554	94	0	-36.553
26	tsuzuku (続く)	verb	225	16	0	-35.120
27	utsukushi (美し)	adjective	178	8	0	-34.680
28	iru (入る)	verb	613	117	10	-34.438

List of Highest Frequency Words for Short Poems in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	sode (袖)	noun	54	11	67	127.217
2	hana (花)	noun	147	47	86	118.196
3	yado (宿)	noun	5	5	37	92.297
4	tsuyu (露)	noun	35	13	47	87.824
5	kimi (君)	pronoun	4	32	42	82.643
6	no (の)	particle (case)	14430	6203	995	80.085
7	ya (や)	particle (bound)	1046	1104	196	70.380
8	aki (秋)	noun	66	20	43	62.444
9	zo (ぞ)	particle (bound)	1182	408	156	62.301
10	kana (かな)	particle (final)	264	318	88	58.056
11	tazunu (訪ぬ)	verb	11	4	25	53.860
12	ushi (憂し)	adjective	109	41	46	51.846
13	o (を)	particle (case)	5952	2856	459	47.665
14	mi (身)	noun	274	272	76	46.147
15	iku (行く)	verb	261	116	62	44.212
16	haru (春)	noun	47	37	33	42.661
17	ama (海女)	noun	13	2	20	40.790
18	ki (き)	auxiliary verb	1524	1395	195	39.498
19	nami (波)	noun	14	7	21	39.480
20	kage (影)	noun	32	13	25	37.999
21	kumoi (雲居)	noun	5	1	16	37.886
22	matsu (松)	noun	21	8	22	37.747
23	iro (色)	noun	138	35	40	37.488
24	domo (ども)	article (conjective	49	18	28	37.344
25	ni (に)	particle (case)	9024	4266	595	33.668
26	kyoo (今日)	noun	107	83	39	33.456
27	kaku (掛く)	verb	176	66	43	32.807
28	na (名)	noun	55	42	28	30.405
29	kiyu (消ゆ)	verb	43	16	23	29.608
30	musubu (結ぶ)	verb	14	3	16	29.514
31	au (会う)	verb	47	31	25	28.906
32	miru (見る)	verb	1848	802	166	28.870
33	sora (空)	noun	146	35	35	28.640
34	koromo (衣)	noun	6	5	14	28.201
35	nuru (濡る)	verb	23	3	17	27.598
36	ura (浦)	noun	12	10	16	27.034
37	yama (山)	noun	77	42	28	26.561
38	kami (神)	noun	27	15	19	26.332
39	kumo (雲)	noun	8	6	14	26.317
40	namida (涙)	noun	142	14	31	25.923
41	tsuki (月)	noun	131	39	32	25.616
42	kaze (風)	noun	103	37	29	25.046
43	mashi (まし)	auxiliary verb	213	140	45	24.844
44	wakaru (別る)	verb	35	20	20	24.815
45	ga (が)	particle (case)	591	313	77	24.708
46	se (瀬)	noun	3	6	12	24.471
47	ominaeshi (女郎花)	noun	6	0	11	24.205
48	ramu (らむ)	auxiliary verb	333	274	60	24.125
49	keburi (煙)	noun	16	8	15	23.923
50	kou (恋ふ)	verb	30	7	17	23.733

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	yo (世)	noun	878	730	110	23.382
52	kusa (草)	noun	14	6	14	23.323
53	mi (み)	suffix	2	3	10	22.414
54	namu (なむ)	particle (final)	15	14	15	22.102
55	uguisu (鶯)	noun	10	2	11	20.119
56	shimo (霜)	noun	6	2	10	20.045
57	fu (経)	verb	118	80	30	19.796
58	shiru (知る)	verb	726	484	86	19.794
59	ne (音)	noun	101	61	27	19.467
60	kagoto (託言)	noun	4	10	11	19.105
61	waku (分く)	verb	133	89	31	18.858
62	oru (折る)	verb	46	11	17	18.829
63	karakoromo (唐衣)	noun	0	1	7	18.811
64	naku (鳴く)	verb	39	4	15	18.192
65	hedatsu (隔つ)	verb	75	50	23	18.063
66	zu (ず)	auxiliary verb	4269	2061	290	17.958
67	tomu (止む)	verb	23	6	13	17.936
68	ne (根)	noun	2	2	8	17.931
69	nobe (野辺)	noun	7	6	10	17.239
70	kazasu (挿頭す)	verb	5	0	8	17.068
71	madou (惑ふ)	verb	107	41	24	16.878
72	kasumi (霞)	noun	16	6	11	15.964
73	tomo (とも)	article (conjective	120	206	35	15.719
74	ka (香)	noun	36	10	14	15.696
75	tayu (絶ゆ)	verb	130	59	26	15.570
76	fuku (吹く)	verb	123	25	23	15.545
77	e (方)	noun	13	11	11	15.337
78	shimizu (清水)	noun	0	0	5	15.288
78	choo (てふ)	verb	0	0	5	15.288
80	kawa (川)	noun	19	6	11	15.041
81	wasuru (忘る)	verb	102	65	24	15.030
82	kurasu (暗す)	verb	13	6	10	14.846
82	tsuyukeshi (露けし)	adjective	13	6	10	14.846
84	mine (峰)	noun	10	4	9	14.513
85	kiri (霧)	noun	20	7	11	14.480
86	ka (か)	particle (bound)	457	616	71	13.992
87	migiwa (汀)	noun	5	1	7	13.766
88	matsu (待つ)	verb	95	45	21	13.719
89	kazashi (挿頭)	noun	9	2	8	13.515
90	michi (道)	noun	109	62	23	13.450
91	sime (注連)	noun	1	2	6	13.448
92	yuki (雪)	noun	59	11	15	13.288
93	nurasu (濡らす)	verb	11	1	8	13.083
94	magau (紛ふ)	verb	21	4	10	13.041
95	fuchi (淵)	noun	0	1	5	12.966
95	kohagi (小萩)	noun	1	0	5	12.966
95	negura (塒)	noun	1	0	5	12.966
98	omooyu (思ほゆ)	verb	14	5	9	12.743
99	tomo (友)	noun	3	5	7	12.610
99	shizuku (滴)	noun	7	1	7	12.610

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
101	ku (来)	verb	332	145	40	12.487
102	musubooru (結ぼおる	verb	11	9	9	12.439
103	yo (夜)	noun	195	69	28	12.374
104	kawaru (変わる)	verb	133	53	23	12.261
105	asu (褪す)	verb	0	0	4	12.230
105	otomego (乙女子)	noun	0	0	4	12.230
105	kusaba (草葉)	noun	0	0	4	12.230
105	sugiushi (過ぎ憂し)	adjective	0	0	4	12.230
105	namiji (波路)	noun	0	0	4	12.230
105	fuetake (笛竹)	noun	0	0	4	12.230
105	matsushima (松島)	noun	0	0	4	12.230
105	moru (守る)	verb	0	0	4	12.230
105	yadorigi (宿り木)	noun	0	0	4	12.230
105	yamazakura (山桜)	noun	0	0	4	12.230
115	chigiru (契る)	verb	38	10	12	11.866
116	wa (わ)	pronoun	343	93	37	11.781
117	sasu (差す)	verb	2	0	5	11.702
117	moshio (藻塩)	noun	2	0	5	11.702
119	yamazato (山里)	noun	36	24	13	11.611
120	nagu (投ぐ)	verb	11	5	8	11.609
120	furusato (古里)	noun	11	5	8	11.609
122	ato (跡)	noun	22	18	11	11.581
123	ma (間)	noun	57	16	14	11.373
124	meguru (巡る)	verb	15	2	8	11.292
125	tou (訪ふ)	verb	1	5	6	11.277
126	karu (枯る)	verb	17	1	8	10.991
127	wakare (別れ)	noun	17	17	10	10.980
128	niou (匂ふ)	verb	47	7	12	10.920
129	aru (荒る)	verb	36	8	11	10.877
130	tama (魂)	noun	2	1	5	10.769
131	ise (伊勢)	noun	7	0	6	10.735
132	murasaki (紫)	noun	34	2	10	10.596
133	suma (須磨)	noun	12	1	7	10.495
134	nagamu (眺む)	verb	173	23	22	10.450
135	utsuru (移る)	verb	39	9	11	10.236
136	haru (晴る)	verb	13	1	7	10.158
136	nagaru (流る)	verb	14	0	7	10.158
138	iwane (岩根)	noun	1	0	4	10.083
138	uranami (浦波)	noun	1	0	4	10.083
138	tazu (田鶴)	noun	1	0	4	10.083
138	tsumu (集む)	verb	1	0	4	10.083
138	tokoyo (常世)	noun	1	0	4	10.083
143	koma (駒)	noun	3	1	5	10.022
144	itsu (何時)	pronoun	58	41	15	9.893
145	kagami (鏡)	noun	12	4	7	9.544
146	tanomu (頼む)	verb	57	75	17	9.483
147	fuji (藤)	noun	21	3	8	9.457
148	magaki (籬)	noun	6	4	6	9.412
149	sato (里)	noun	29	14	10	9.403
150	karu (刈る)	verb	4	1	5	9.397

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
150	shima (島)	noun	4	1	5	9.397
150	hototogisu (時鳥)	noun	5	0	5	9.397
153	moga (もが)	particle (final)	14	11	8	9.238
154	chiru (散る)	verb	57	10	12	9.210
155	akigiri (秋霧)	noun	0	0	3	9.172
155	asajuu (浅茅生)	noun	0	0	3	9.172
155	amabune (海人舟)	noun	0	0	3	9.172
155	ayame (菖蒲)	noun	0	0	3	9.172
155	oosaka (あふさか)	noun	0	0	3	9.172
155	ogushi (小櫛)	noun	0	0	3	9.172
155	kozutau (木伝う)	verb	0	0	3	9.172
155	sayu (冴ゆ)	verb	0	0	3	9.172
155	tabigoromo (旅衣)	noun	0	0	3	9.172
155	tamakazura (玉葛)	noun	0	0	3	9.172
155	natsugoromo (夏衣)	noun	0	0	3	9.172
155	fujinami (藤波)	noun	0	0	3	9.172
155	funabito (船人)	noun	0	0	3	9.172
155	furukawa (古川)	noun	0	0	3	9.172
155	miyama (深山)	noun	0	0	3	9.172
155	yaso (八十)	noun	0	0	3	9.172
155	yorozuyo (万代)	noun	0	0	3	9.172
155	rashi (らし)	auxiliary verb	0	0	3	9.172
173	kanashi (悲し)	adjective	219	88	27	9.106
174	kakine (垣根)	noun	8	3	6	9.046
175	ume (梅)	noun	22	4	8	9.027
176	sugi (杉)	noun	1	1	4	8.965
176	aoi (葵)	noun	2	0	4	8.965
178	tatsu (断つ)	verb	3	3	5	8.861
178	tamoto (袂)	noun	4	2	5	8.861
178	mayou (迷う)	verb	5	1	5	8.861
178	seki (関)	noun	6	0	5	8.861
182	ayanashi (文無し)	adjective	2	10	6	8.707
183	ji (じ)	auxiliary verb	226	289	37	8.686
184	na (な)	particle (final)	34	126	18	8.573
185	saku (咲く)	verb	28	10	9	8.573
186	nadeshiko (撫子)	noun	16	4	7	8.504
186	shiotaru (潮垂る)	verb	20	0	7	8.504
188	furu (降る)	verb	51	10	11	8.491
189	kishi (岸)	noun	6	1	5	8.391
190	shi (し)	particle (adverbial	50	13	11	8.259
191	uji (宇治)	noun	22	8	8	8.257
192	sono (園)	noun	1	2	4	8.156
192	shio (潮)	noun	2	1	4	8.156
194	katami (形見)	noun	25	16	9	8.116
195	fumu (踏む)	verb	10	4	6	8.099
195	kiku (菊)	noun	11	3	6	8.099
197	kutsu (朽つ)	verb	5	3	5	7.973
197	akegure (明け暗れ)	noun	8	0	5	7.973
197	tachibana (橘)	noun	8	0	5	7.973

List of Lowest Frequency Words for Short Poems in *The Tale of Genji*

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	tamau (給ふ: respect)	verb	11207	2691	0	-368.635
2	oon (御)	prefix	3696	1086	0	-125.499
3	ito (いと)	adverb	3044	1178	0	-110.730
4	nado (など)	particle (adverbial)	3229	608	0	-100.588
5	obosu (思す)	verb	2782	584	0	-88.193
6	kikoyu (聞こゆ)	verb	2691	961	3	-83.750
7	nari (なり: declarative)	auxiliary verb	10587	4800	218	-73.477
8	haberi (侍り)	verb	7	2408	0	-63.206
9	koto (事)	noun	2754	1931	33	-52.754
10	tari (たり: perfective)	auxiliary verb	3379	909	31	-47.259
11	namu (なむ)	particle (bound)	359	1416	0	-46.422
12	mi (御)	prefix	1616	578	4	-44.230
13	tatematsuru (奉る)	verb	1167	504	0	-43.697
14	sama (様)	noun	1351	470	1	-43.292
15	domo (共)	suffix	1129	235	0	-35.656
16	te (て)	particle (conjective)	11563	3414	283	-33.437
17	ge (気)	suffix	1118	259	1	-31.904
18	notamau (宣ふ)	verb	1065	139	0	-31.468
19	sari (然り)	verb	611	493	0	-28.851
20	ari (有り)	verb	2924	1658	56	-27.852
21	ni (に)	particle (conjective)	1691	388	12	-26.854
22	sasu (さす)	auxiliary verb	541	464	0	-26.261
23	mairu (参る)	verb	733	257	0	-25.868
24	miya (宮)	noun	782	183	0	-25.214
25	su (す)	auxiliary verb	804	490	3	-24.571
26	yoo (様)	noun	496	349	0	-22.076
27	owasu (おはす)	verb	645	145	0	-20.638
28	hitobito (人々)	noun	604	107	0	-18.572
29	arisama (有り様)	noun	483	216	0	-18.259
30	imiji (いみじ)	adjective	543	155	0	-18.232
31	okashi (をかし)	adjective	589	73	0	-17.291
32	gimi (君)	suffix	554	94	0	-16.925
33	ka (彼)	pronoun	508	247	1	-16.153
34	monosu (ものす)	verb	207	338	0	-14.233
35	rei (例)	noun	449	79	0	-13.789
36	ooshi (多し)	adjective	379	142	0	-13.606
37	otodo (大臣)	noun	413	101	0	-13.423
38	ayashi (怪し)	adjective	296	217	0	-13.397
39	owashimasu (おはします)	verb	288	220	0	-13.267
40	izu (出ず)	verb	1484	358	20	-13.230
41	tamau (給ふ: humble)	verb	3	498	0	-13.084
42	kashi (かし)	particle (final)	337	264	1	-12.326
43	nasu (なす)	verb	553	292	4	-12.218
44	ri (り)	auxiliary verb	2886	501	54	-12.180
45	oboyu (覚ゆ)	verb	530	237	3	-12.147
46	meri (めり)	auxiliary verb	472	504	6	-12.073
47	ba (ば)	particle (conjective)	3966	1242	97	-11.984
48	maji (まじ)	auxiliary verb	287	286	1	-11.636
49	sukoshi (少し)	adverb	432	127	1	-11.291
50	onna (女)	noun	307	121	0	-11.176

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	mote (以て)	prefix	444	195	2	-11.045
52	do (ど)	particle (conjective)	1861	634	37	-10.580
53	in (院)	noun	311	91	0	-10.497
54	kaku (斯く)	adverb	606	526	10	-10.337
55	itooshi (いとおし)	adjective	284	102	0	-10.079
56	saburau (侍ふ)	verb	250	134	0	-10.027
57	nari (なり: hearsay)	auxiliary verb	115	262	0	-9.844
58	tachi (達)	suffix	290	82	0	-9.713
59	sa (然)	adverb	350	302	3	-9.564
60	kehai (気配)	noun	331	35	0	-9.557
61	raru (らる)	auxiliary verb	357	284	3	-9.321
62	koo (斯う)	adverb	160	174	0	-8.721
63	hodo (程)	noun	1163	541	23	-8.701
64	kokorogurushi (心苦し)	adjective	217	111	0	-8.564
65	waza (技)	noun	153	170	0	-8.433
66	moosu (申す)	verb	201	118	0	-8.329
67	kaku (書く)	verb	377	55	1	-8.197
68	yorozu (万)	noun	222	91	0	-8.172
69	kuchioshi (口惜し)	adjective	197	98	0	-7.702
70	beshi (べし)	auxiliary verb	1719	1404	57	-7.657
71	to (と)	particle (case)	7816	1973	229	-7.619
72	zu (ず)	verb	171	116	0	-7.493
73	taishoo (大将)	noun	226	58	0	-7.415
74	mashite (まして)	adverb	199	84	0	-7.389
75	tada (唯)	adverb	396	218	4	-7.287
76	omou (思ふ)	verb	2867	1639	92	-7.270
77	toshigoro (年頃)	noun	165	112	0	-7.232

List of Highest Frequency Words for Narration in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	tari (たり)	auxiliary verb	3997	1205	62	545.485
2	iu (言ふ)	verb	3244	717	168	534.322
3	nado (など)	particle (adverbial)	2437	540	1	495.338
4	te (て)	particle (conjective)	9482	4002	884	491.982
5	to (と)	particle (case)	6789	2846	1093	232.090
6	okashi (をかし)	adjective	721	105	5	194.204
7	otoko (男)	noun	661	100	0	178.960
8	ba (ば)	particle (conjective)	4505	1803	738	169.911
9	ni (に)	particle (conjective)	1292	331	89	168.706
10	tote (とて)	particle (case)	1253	365	54	156.951
11	keri (けり)	auxiliary verb	3569	1519	530	125.503
12	onna (女)	noun	600	134	0	120.656
13	uchi (打ち)	prefix	632	119	48	107.110
14	kaku (書く)	verb	427	78	10	92.172
15	iru (入る)	verb	633	161	30	91.371
16	ge (気)	suffix	604	178	10	83.634
17	iru (居る)	verb	504	106	30	83.384
18	hodo (程)	noun	962	368	52	74.130
19	yaru (遣る)	verb	447	84	40	71.600
20	warau (笑ふ)	verb	275	49	0	66.570
21	ito (いと)	adverb	1910	1092	13	64.776
22	domo (共)	suffix	699	279	2	63.871
23	kaeshi (返し)	noun	199	22	3	59.877
24	yomu (読む)	verb	365	106	4	52.727
25	fusu (伏す)	verb	226	37	9	49.273
26	iku (行く)	verb	242	59	1	43.877
27	ka (日)	suffix	250	58	7	43.087
28	kaerigoto (返り事)	noun	170	29	0	42.448
29	moto (下)	noun	324	94	11	42.366
30	hiku (引く)	verb	328	82	27	41.347
31	shiroshi (白し)	adjective	147	20	3	38.905
32	nonoshiru (罵る)	verb	134	19	0	37.685
33	izu (出ず)	verb	938	386	134	37.134
34	rei (例)	noun	257	79	0	36.153
35	garu (がる)	suffix	154	29	0	35.744
36	sashi (差し)	prefix	240	63	8	35.498
37	okosu (遣す)	verb	128	17	3	33.905
38	tsutomete (つとめて)	noun	94	8	0	33.620
39	fumi (文)	noun	314	103	13	33.434
40	uta (歌)	noun	263	85	2	33.372
41	kokochi (心地)	noun	349	124	14	32.878
42	idasu (出だす)	verb	183	47	0	31.957
43	naku (泣く)	verb	229	33	38	31.841
44	shooshoo (少将)	noun	213	63	0	31.510
45	aru (ある)	attributive word	106	15	0	29.844
46	hitobito (人々)	noun	357	149	2	29.398
47	tachiwaki (帯刀)	noun	80	9	0	25.432
48	notamau (宣ふ)	verb	504	260	0	25.193
49	ayumu (歩む)	verb	60	3	0	25.172
50	tsuku (付く)	verb	452	200	25	25.064

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	kinu (衣)	noun	122	26	3	23.079
52	tokoro (所)	noun	542	288	7	22.829
53	warawa (童)	noun	112	23	2	22.503
54	ari (あり)	verb	2779	1620	435	22.455
55	oboyu (覚ゆ)	verb	459	237	2	22.298
56	kami (髪)	noun	75	7	3	21.845
57	tsugomori (晦)	noun	52	3	0	21.040
58	monogatari (物語)	noun	117	30	0	20.464
59	yobu (呼ぶ)	verb	121	31	1	20.457
60	kiku (聞く)	verb	699	284	135	20.404
61	tsuitachi (一日)	noun	56	5	0	19.665
62	miyu (見ゆ)	verb	553	183	133	19.565
63	kichoo (几帳)	noun	89	18	0	19.501
64	kiru (着る)	verb	146	26	21	19.263
65	ariku (歩く)	verb	140	43	1	19.103
66	mina (皆)	noun	247	102	6	18.923
67	hiru (昼)	noun	77	13	1	18.353
68	hashiru (走る)	verb	72	11	1	18.266
69	hashi (端)	noun	69	9	2	18.036
70	agu (上ぐ)	verb	123	30	7	17.798
71	myoobu (命婦)	noun	41	2	0	17.301
72	akogi (あこぎ)	noun	92	22	0	17.299
73	sawagu (騒ぐ)	verb	125	22	17	17.227
74	kiyoshi (清し)	adjective	104	21	8	16.540
75	aku (明く)	verb	226	65	36	16.477
76	yooyoo (漸う)	adverb	57	8	0	16.123
77	naishi (内侍)	noun	59	9	0	15.867
78	omou (思ふ)	verb	1892	876	517	15.842
79	orimono (織物)	noun	45	4	0	15.829

List of Lowest Frequency Words for Narration in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	haberi (侍り)	verb	209	1284	2	-421.910
2	mu (む)	auxiliary verb	1234	2219	604	-350.495
3	ki (き)	auxiliary verb	979	1970	470	-346.404
4	ha (は)	particle (bound)	3593	4179	1591	-328.219
5	koso (こそ)	particle (bound)	522	1174	282	-240.898
6	ya (や)	particle (bound)	597	978	580	-235.721
7	kana (かな)	particle (final)	69	264	206	-162.617
8	ka (か)	particle (bound)	368	691	315	-161.869
9	kimi (君)	pronoun	1	34	224	-161.148
10	tamau (給ふ: respect)	verb	2779	3915	3	-140.288
11	owashimasu (おはします)	verb	193	664	0	-139.405
12	oon (御)	prefix	995	1673	2	-107.347
13	moosu (申す)	verb	322	776	2	-104.746
14	zo (ぞ)	particle (bound)	1064	1119	622	-102.982
15	daijin (大臣)	noun	13	211	0	-100.174
16	ima (今)	noun	236	456	130	-82.205
17	ramu (らむ)	auxiliary verb	194	224	295	-80.984
18	ji (じ)	auxiliary verb	123	252	150	-79.923
19	sasu (さす)	auxiliary verb	390	795	1	-79.696
20	dono (殿)	suffix	147	438	0	-78.412
21	nado (など)	adverb	26	179	26	-76.116
22	tomo (とも)	particle (conjective)	49	156	96	-73.767
23	tamau (給ふ: humble)	verb	8	144	0	-70.058
24	na (な)	particle (final)	32	135	71	-69.113
25	nyuudoo (入道)	noun	3	122	0	-68.001
26	nen (年)	noun	12	152	0	-67.395
27	koto (事)	noun	1407	1792	174	-67.279
28	na (な)	adverb	5	103	25	-66.775
29	tsu (つ)	auxiliary verb	606	818	191	-62.549
30	so (そ)	particle (final)	6	99	25	-62.149
31	su (す)	auxiliary verb	893	1320	21	-60.233
32	ga (が)	particle (case)	600	554	427	-57.718
33	mi (身)	noun	114	112	218	-57.015
34	in (院)	noun	58	234	0	-56.936
35	tatematsuru (奉る)	verb	343	642	4	-54.538
36	mashi (まし)	auxiliary verb	63	114	122	-53.745
37	yo (世)	noun	250	326	193	-53.665
38	namu (なむ)	particle (bound)	539	885	2	-53.158
39	yado (宿)	noun	2	3	84	-48.866
40	ana (あな)	interjection	24	135	13	-48.470
41	daijoo (太政)	noun	1	81	0	-48.058
42	toshi (年)	noun	91	182	84	-46.503
43	aki (秋)	noun	45	17	169	-46.188
44	owasu (おはす)	verb	194	411	0	-44.153
45	nari (なり: hearsay)	auxiliary verb	132	269	51	-43.239
46	wa (わ)	pronoun	190	131	266	-41.545
47	maro (麻呂)	pronoun	0	61	3	-41.233
48	kashi (かし)	particle (final)	176	362	13	-40.712
49	ichi (一)	noun	57	195	0	-40.646
50	tennoo (天皇)	noun	1	69	0	-40.464

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	ikani (如何に)	adverb	140	278	43	-39.653
52	o (を)	particle (conjective)	45	85	84	-38.570
53	mi (み)	suffix	1	1	65	-38.569
54	beshi (べし)	auxiliary verb	753	874	188	-37.615
55	uketamawaru (承る)	verb	10	94	0	-37.586
56	haya (早)	adverb	3	66	6	-37.132
57	nani (何)	pronoun	143	228	89	-36.959
58	hana (花)	noun	118	43	236	-36.238
59	otodo (大臣)	noun	128	293	0	-36.090
60	iza (いざ)	interjection	1	39	23	-36.046
61	ushi (憂し)	adjective	34	30	113	-36.022
62	kurai (位)	noun	9	87	0	-35.165
63	zu (ず)	auxiliary verb	2233	1636	987	-34.824
64	ware (我)	pronoun	192	178	197	-34.248
65	yo (よ)	particle (final)	45	147	12	-34.232
66	oboshimesu (思し召す)	verb	45	159	0	-34.232
67	koko (此処)	pronoun	73	192	13	-34.074
68	makaru (罷る)	verb	25	121	0	-33.945
69	kanpaku (関白)	noun	5	73	0	-33.709
70	kou (恋ふ)	verb	13	5	88	-32.957
71	namu (なむ)	particle (final)	11	12	76	-32.880
72	oba (をば)	particle (case)	83	166	51	-32.767
73	haru (春)	noun	36	16	123	-32.532
74	saburau (侍ふ)	verb	158	322	0	-32.087
75	ta (誰)	pronoun	7	56	20	-32.045
76	sa (然)	adverb	221	371	32	-31.539
77	dani (だに)	particle (adverbial)	156	197	118	-30.814
78	kyoo (今日)	noun	83	112	98	-30.252
79	ikaga (如何)	adverb	90	203	16	-29.753
80	sari (然り)	verb	319	508	7	-29.133
81	koishi (恋し)	adjective	28	21	95	-28.855
82	tanomu (頼む)	verb	30	31	87	-28.075
83	tadaima (只今)	noun	30	118	0	-28.075
84	guu (宮)	suffix	2	53	0	-27.806
85	mi (御)	prefix	385	563	22	-27.280
86	ide (いで)	interjection	8	67	3	-27.189
87	mikado (帝)	noun	108	233	0	-25.870
88	shiru (知る)	verb	374	312	251	-25.421
89	usu (失す)	verb	49	143	1	-25.330
90	sode (袖)	noun	39	13	114	-25.129
91	choo (てふ)	verb	2	1	47	-24.754
92	yama (山)	noun	131	63	197	-24.548
93	zoo (贈)	noun	0	38	0	-24.481
94	koi (恋)	noun	3	1	50	-24.478
95	chiru (散る)	verb	35	8	110	-24.222
96	kaku (斯く)	adjective	321	449	40	-22.993
97	sakurabana (桜花)	noun	0	0	35	-22.548
98	tamawaru (賜わる)	verb	11	68	0	-22.296
99	omoi (思ひ)	noun	13	11	61	-22.108
100	onore (己)	pronoun	4	48	2	-22.075

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
101	nageki (嘆き)	noun	1	14	25	-21.625
102	nushi (主)	noun	3	36	10	-21.518
103	no (の)	particle (case)	11169	7793	3338	-21.391
104	naku (鳴く)	verb	62	8	145	-21.232
105	tooguu (東宮)	noun	8	59	0	-21.221
106	tsuyu (露)	noun	21	11	74	-20.730
107	sanjoo (三条)	noun	0	32	0	-20.615
108	suke (出家)	noun	0	32	0	-20.615
109	reizei (冷泉)	noun	0	32	0	-20.615
110	tame (為)	noun	18	48	31	-20.591
111	ikade (如何で)	adverb	121	207	23	-19.791
112	ama (海女)	noun	3	2	41	-19.755
113	yoso (余所)	noun	8	18	38	-19.628
114	kore (此れ)	pronoun	270	389	21	-19.065
115	au (会ふ)	verb	129	74	163	-18.867
116	umareru (生まる)	verb	2	37	1	-18.707
117	akikaze (秋風)	noun	0	2	27	-18.683
118	kyoo (卿)	suffix	27	91	0	-18.672
119	inochi (命)	noun	15	37	32	-18.646
120	ko (是)	pronoun	848	1000	41	-18.450
121	meri (めり)	auxiliary verb	289	407	22	-18.426
122	rashi (らし)	auxiliary verb	0	1	27	-18.038
123	izura (いづら)	pronoun	0	21	7	-18.038
124	nai (内)	prefix	0	28	0	-18.038
125	koogoo (皇后)	noun	1	33	0	-17.903
126	shi (し)	particle (adverbial)	56	21	113	-17.730
127	tose (年)	suffix	25	65	20	-17.600
128	domo (ども)	particle (conjective)	90	80	101	-17.553
129	kashikoshi (賢し)	adjective	36	97	4	-16.762
130	beranari (べらなり)	auxiliary verb	0	1	25	-16.750
131	momijiba (紅葉)	noun	1	1	30	-16.668
132	dai (第)	prefix	1	31	0	-16.668
133	too (東)	noun	1	31	0	-16.668
134	sue (末)	noun	32	74	20	-16.523
135	obosu (思す)	verb	239	360	0	-16.276
136	tama (玉)	noun	13	29	31	-16.254

List of Highest Frequency Words for Conversation in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	haberi (侍り)	verb	209	1284	2	698.395
2	tamau (給ふ: respect)	verb	2779	3915	3	621.402
3	oon (御)	prefix	995	1673	2	344.100
4	ki (き)	auxiliary verb	979	1970	470	291.710
5	owashimasu (おはします)	verb	193	664	0	269.047
6	mu (む)	auxiliary verb	1234	2219	604	260.526
7	moosu (申す)	verb	322	776	2	237.304
8	su (す)	auxiliary verb	893	1320	21	216.290
9	sasu (さす)	auxiliary verb	390	795	1	206.148
10	koso (こそ)	particle (bound)	522	1174	282	195.755
11	koto (事)	noun	1407	1792	174	182.055
12	namu (なむ)	particle (bound)	539	885	2	175.112
13	dono (殿)	suffix	147	438	0	160.396
14	daijin (大臣)	noun	13	211	0	149.796
15	tatematsuru (奉る)	verb	343	642	4	149.154
16	ha (は)	bound (bound)	3593	4179	1591	134.391
17	owasu (おはす)	verb	194	411	0	111.218
18	in (院)	noun	58	234	0	104.311
19	tamau (給ふ: humble)	verb	8	144	0	104.069
20	nen (年)	noun	12	152	0	102.672
21	ko (是)	pronoun	848	1000	41	99.364
22	nyuudoo (入道)	noun	3	122	0	97.478
23	sari (然り)	verb	319	508	7	93.502
24	kashi (かし)	particle (final)	176	362	13	87.562
25	otodo (大臣)	noun	128	293	0	85.688
26	mi (御)	prefix	385	563	22	85.420
27	saburau (侍ふ)	verb	158	322	0	83.553
28	ichi (一)	noun	57	195	0	78.597
29	nado (など)	adverb	26	179	26	72.437
30	daijoo (太政)	noun	1	81	0	67.801
31	oboshimesu (思し召す)	verb	45	159	0	65.432
32	mikado (帝)	noun	108	233	0	64.210
33	tsu (つ)	auxiliary verb	606	818	191	63.878
34	obosu (思す)	verb	239	360	0	62.858
35	sa (然)	adverb	221	371	32	62.067
36	mairu (参る)	verb	474	554	1	59.783
37	makaru (罷る)	verb	25	121	0	59.310
38	uketamawaru (承る)	verb	10	94	0	58.945
39	tennoo (天皇)	noun	1	69	0	57.255
40	ima (今)	noun	236	456	130	56.707
41	ana (あな)	interjection	24	135	13	56.683
42	meri (めり)	auxiliary verb	289	407	22	55.757
43	kaku (斯く)	adverb	321	449	40	55.639
44	kore (此れ)	pronoun	270	389	21	55.449
45	kurai (位)	noun	9	87	0	54.978
46	koko (此処)	pronoun	73	192	13	54.769
47	kikoyu (聞こゆ)	verb	471	548	19	53.572
48	tadaima (只今)	noun	30	118	0	51.823
49	ka (か)	particle (bound)	368	691	315	51.735
50	kanpaku (関白)	noun	5	73	0	50.781

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	usu (失す)	verb	49	143	1	50.707
52	beshi (べし)	auxiliary verb	753	874	188	49.836
53	ikani (如何に)	adverb	140	278	43	49.170
54	ikaga (如何)	adverb	90	203	16	49.064
55	yo (よ)	particle (final)	45	147	12	47.904
56	nari (なり: hearsay)	auxiliary verb	132	269	51	45.159
57	maro (麻呂)	pronoun	0	61	3	44.939
58	na (な)	adverb	5	103	25	41.600
59	guu (宮)	suffix	2	53	0	40.487
60	haya (早)	adverb	3	66	6	38.322
61	so (そ)	particle (final)	6	99	25	38.072
62	tamawaru (賜わる)	verb	11	68	0	37.084
63	ide (いで)	interjection	8	67	3	36.325
64	kyoo (卿)	suffix	27	91	0	36.306
65	ya (や)	particle (bound)	597	978	580	35.714
66	tono (殿)	noun	199	260	1	35.144
67	tooguu (東宮)	noun	8	59	0	34.323
68	zoo (贈)	noun	0	38	0	33.824
69	ikade (如何で)	adverb	121	207	23	33.490
70	nari (なり: declarative)	auxiliary verb	5409	4119	771	33.098
71	kashikoshi (賢し)	adjective	36	97	4	29.928
72	onore (己)	pronoun	4	48	2	28.681
73	sanjoo (三条)	noun	0	32	0	28.483
73	suke (出家)	noun	0	32	0	28.483
73	reizei (冷泉)	noun	0	32	0	28.483
76	ni (二)	noun	36	89	0	27.942
77	hara (腹)	noun	23	72	0	27.299
78	koogoo (皇后)	noun	1	33	0	25.841
79	ooji (祖父)	noun	5	42	0	25.481
80	nai (内)	prefix	0	28	0	24.922
81	umaru (生まる)	verb	2	37	1	24.837
82	shimu (しむ)	auxiliary verb	4	39	0	24.700
83	dai (第)	prefix	1	31	0	24.114
83	too (東)	noun	1	31	0	24.114
85	so (其)	prefix	418	407	18	23.558
86	nyoogo (女御)	noun	23	66	0	23.457
87	miko (御子)	noun	86	131	0	23.253
88	kisaki (后)	noun	16	56	0	22.898
89	haha (母)	noun	45	90	0	22.851
90	binnashi (便無し)	adjective	14	53	0	22.753
91	kookoo (斯う斯う)	adverb	5	38	0	22.329
92	ichijoo (一条)	noun	0	25	0	22.252
92	koo (公)	suffix	0	25	0	22.252
92	moosu (申す)	verb	0	25	0	22.252
95	goran (御覧)	noun	86	128	0	21.909
96	sesshoo (摂政)	noun	1	28	0	21.530
97	sore (其れ)	pronoun	213	239	9	21.194
98	ashi (悪し)	adjective	71	114	2	21.015
99	nanigoto (何事)	noun	48	92	3	20.668
100	kazan (花山)	noun	0	23	0	20.472

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
100	kootaigoo (皇太后)	noun	0	23	0	20.472
102	kyoo (興)	noun	7	39	0	20.378
103	oba (をば)	particle (case)	83	166	51	20.374
104	san (三)	noun	56	96	0	20.189
105	chuuguu (中宮)	noun	8	40	0	19.908
106	murakami (村上)	noun	1	26	0	19.813
107	toshi (疾し)	adjective	60	115	17	19.791
108	tachi (達)	suffix	156	182	0	19.618
109	gannen (元年)	noun	0	22	0	19.582
110	zu (ず)	verb	97	131	0	19.004
111	tsukoomatsuru (仕る)	verb	77	113	0	18.907
112	na (な)	particle (final)	32	135	71	18.525
113	hon (品)	suffix	4	31	0	18.333
114	ori (折)	noun	145	178	13	17.680
115	waka (和歌)	noun	1	23	0	17.247
116	sanjoo (三条)	noun	12	42	0	17.173
117	kochi (此方)	pronoun	0	19	0	16.911
118	kana (かな)	particle (final)	69	264	206	16.773
119	ichijoo (一条)	noun	19	50	0	16.570
120	muzu (むず)	auxiliary verb	6	32	0	16.404
121	a (吾)	pronoun	1	22	0	16.395
122	genpuku (元服)	noun	0	18	0	16.021
122	tsukamatsuru (仕る)	verb	0	18	0	16.021
122	doo (同)	noun	0	18	0	16.021
125	yoshi (良し)	adjective	216	223	8	15.876
126	juu (十)	noun	131	151	0	15.829
127	kanarazu (必ず)	adverb	35	66	0	15.654
128	nani (何)	pronoun	143	228	89	15.603

List of Lowest Frequency Words for Conversation in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	iu (言ふ)	verb	3244	717	168	-302.934
2	nado (など)	particle (adverbial)	2437	540	1	-198.867
3	to (と)	particle (case)	6789	2846	1093	-190.894
4	te (て)	particle (conjective)	9482	4002	884	-188.690
5	tari (たり)	auxiliary verb	3997	1205	62	-182.457
6	ba (ば)	particle (conjective)	4505	1803	738	-151.317
7	okashi (をかし)	adjective	721	105	5	-100.134
8	ni (に)	conjective (conjective)	1292	331	89	-98.091
9	keri (けり)	auxiliary verb	3569	1519	530	-88.015
10	otoko (男)	noun	661	100	0	-87.235
11	uchi (打ち)	prefix	632	119	48	-76.558
12	tsutsu (つつ)	particle (conjective)	315	57	143	-72.031
13	tote (とて)	particle (case)	1253	365	54	-67.750
14	naku (鳴く)	verb	62	8	145	-57.390
15	hana (花)	noun	118	43	236	-56.751
16	yaru (遣る)	verb	447	84	40	-55.745
17	omou (思ふ)	verb	1892	876	517	-55.683
18	iku (行く)	verb	269	81	201	-53.849
19	iru (居る)	verb	504	106	30	-51.093
20	onna (女)	noun	600	134	0	-48.043
21	kaku (書く)	verb	427	78	10	-48.014
22	iru (入る)	verb	633	161	30	-45.881
23	aki (秋)	noun	45	17	169	-45.053
24	namida (涙)	noun	64	5	87	-43.558
25	naku (泣く)	verb	229	33	38	-42.208
26	miyu (見ゆ)	verb	553	183	133	-39.337
27	miru (見る)	verb	1299	644	457	-39.102
28	mo (も)	particle (bound)	4506	2584	1227	-38.081
29	chiru (散る)	verb	35	8	110	-35.854
30	kaeshi (返し)	noun	199	22	3	-35.118
31	koe (声)	noun	167	37	81	-33.152
32	yama (山)	noun	131	63	197	-32.808
33	sode (袖)	noun	39	13	114	-31.065
34	warau (笑ふ)	verb	275	49	0	-30.266
35	furu (降る)	verb	162	46	103	-30.091
36	fusu (伏す)	verb	226	37	9	-29.739
37	kimi (君)	pronoun	1	34	224	-29.697
38	haru (春)	noun	36	16	123	-29.163
39	ku (来)	verb	582	267	246	-29.152
40	ge (気)	suffix	604	178	10	-29.021
41	kaze (風)	noun	86	28	114	-28.401
42	nagamu (眺む)	verb	87	8	35	-28.135
43	yo (夜)	noun	323	106	101	-27.747
44	oto (音)	noun	118	16	35	-27.354
45	fuku (吹く)	verb	95	26	93	-26.996
46	hiku (引く)	verb	328	82	27	-26.842
47	ni (に)	particle (case)	7874	4767	1873	-26.724
48	shi (し)	particle (adverbial)	56	21	113	-26.595
49	hito (人)	noun	2061	1072	466	-26.277
50	kou (恋ふ)	verb	13	5	88	-25.955

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	izu (出ず)	verb	938	386	134	-25.671
52	iro (色)	noun	92	32	110	-25.396
53	fu (経)	verb	80	27	104	-24.995
54	kawa (川)	noun	56	12	73	-24.855
55	kiku (聞く)	verb	699	284	135	-24.565
56	o (を)	particle (case)	3647	2178	1062	-24.527
57	yado (宿)	noun	2	3	84	-24.489
58	mizu (水)	noun	67	17	77	-23.591
59	hodo (程)	noun	962	368	52	-23.520
60	tatsu (立つ)	verb	661	287	169	-23.156
61	shiroshi (白し)	adjective	147	20	3	-22.248
62	wa (わ)	pronoun	190	131	266	-22.033
63	mi (み)	suffix	1	1	65	-21.865
64	kiru (着る)	verb	146	26	21	-21.379
65	kiyu (消ゆ)	verb	26	8	75	-21.290
66	kurasu (暮らす)	verb	54	3	22	-20.943
67	ka (日)	suffix	250	58	7	-20.207
68	saku (咲く)	verb	40	11	69	-19.953
69	okosu (遣す)	verb	128	17	3	-19.886
70	kaerigoto (返り事)	noun	170	29	0	-19.695
71	tsutomete (つとめて)	noun	94	8	0	-19.066
72	nu (寝)	verb	118	38	76	-18.890
73	nonoshiru (罵る)	verb	134	19	0	-18.803
74	nami (波)	noun	32	8	61	-18.751
75	au (会ふ)	verb	129	74	163	-18.598
76	ori (居り)	verb	65	6	18	-18.308
77	sawagu (騒ぐ)	verb	125	22	17	-18.269
78	yomu (読む)	verb	365	106	4	-17.827
79	tayu (絶ゆ)	verb	84	36	98	-17.483
80	aku (明く)	verb	226	65	36	-17.393
81	sora (空)	noun	70	21	65	-17.296
82	moto (下)	noun	324	94	11	-17.123
83	koi (恋)	noun	3	1	50	-17.033
84	nashi (無し)	adjective	865	483	333	-16.890
85	hototogisu (時鳥)	noun	25	8	62	-16.878
86	iku (行く)	verb	242	59	1	-16.801
87	kaeru (返る)	verb	337	133	91	-16.797
88	kumo (雲)	noun	27	5	46	-16.510
89	omooyu (思ほゆ)	verb	7	1	44	-16.293
90	niou (匂ふ)	verb	15	0	27	-16.220
91	matsu (松)	noun	36	6	40	-16.047
92	akasu (明かす)	verb	72	10	20	-16.013
93	garu (がる)	suffi	154	29	0	-15.804
94	tsuyu (露)	noun	21	11	74	-15.792
95	nagaru (流る)	verb	18	6	57	-15.727
96	sashi (差し)	prefix	240	63	8	-15.719
97	choo (てふ)	verb	2	1	47	-15.555
98	uguisu (鶯)	noun	15	2	40	-15.495
99	ayumu (歩む)	verb	60	3	0	-15.365
100	kami (髪)	noun	75	7	3	-15.350
101	aru (ある)	attributive word	106	15	0	-14.901

List of Highest Frequency Words for Short Poem in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	kimi (君)	pronoun	1	34	224	331.853
2	hana (花)	noun	118	43	236	224.099
3	aki (秋)	noun	45	17	169	204.659
4	wa (わ)	pronoun	190	131	266	178.964
5	ramu (らむ)	auxiliary verb	194	224	295	174.664
6	mi (身)	noun	114	112	218	163.149
7	ya (や)	particle (bound)	597	978	580	160.955
8	naku (鳴く)	verb	62	8	145	159.682
9	yama (山)	noun	131	63	197	152.414
10	haru (春)	noun	36	16	123	142.113
11	yado (宿)	noun	2	3	84	140.160
12	chiru (散る)	verb	35	8	110	130.480
13	zu (ず)	auxiliary verb	2233	1636	987	128.414
14	sode (袖)	noun	39	13	114	128.197
15	kou (恋ふ)	verb	13	5	88	123.792
16	ga (が)	particle (case)	600	554	427	119.188
17	ushi (憂し)	adjective	34	30	113	116.676
18	mi (み)	suffix	1	1	65	113.374
19	kana (かな)	particle (final)	69	264	206	108.217
20	shi (し)	particle (adverbial)	56	21	113	107.268
21	au (会ふ)	verb	129	74	163	106.988
22	zo (ぞ)	particle (bound)	1064	1119	622	105.865
23	koishi (恋し)	adjective	28	21	95	101.928
24	ha (は)	particle (bound)	3593	4179	1591	99.318
25	iku (行く)	verb	269	81	201	98.115
26	namu (なむ)	particle (final)	11	12	76	97.326
27	no (の)	particle (case)	11169	7793	3338	96.270
28	ware (我)	pronoun	192	178	197	88.768
29	kaze (風)	noun	86	28	114	87.277
30	tsuyu (露)	noun	21	11	74	84.801
31	kiyu (消ゆ)	verb	26	8	75	84.509
32	tanomu (頼む)	verb	30	31	87	81.423
33	koi (恋)	noun	3	1	50	81.095
34	fu (経)	verb	80	27	104	78.152
35	kokoro (心)	noun	462	312	284	78.072
36	choo (てふ)	verb	2	1	47	77.931
37	iro (色)	noun	92	32	110	77.663
38	namida (涙)	noun	64	5	87	76.383
39	omoi (思ひ)	noun	13	11	61	72.182
40	mashi (まし)	auxiliary verb	63	114	122	70.668
41	shiru (知る)	verb	374	312	251	68.657
42	nagaru (流る)	verb	18	6	57	65.922
43	hototogisu (時鳥)	noun	25	8	62	65.671
44	matsu (待つ)	verb	93	47	105	65.278
45	sakurabana (桜花)	noun	0	0	35	65.128
46	tayu (絶ゆ)	verb	84	36	98	65.122
47	omooyu (思ほゆ)	verb	7	1	44	63.355
48	ama (海女)	noun	3	2	41	63.101
49	saku (咲く)	verb	40	11	69	62.870
50	fuku (吹く)	verb	95	26	93	58.959

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	nami (波)	noun	32	8	61	58.959
52	ka (か)	particle (bound)	368	691	315	58.682
53	kawa (川)	noun	56	12	73	58.415
54	ne (音)	noun	19	9	54	57.800
55	mizu (水)	noun	67	17	77	55.609
56	nomi (のみ)	particle (adverbial)	248	131	161	54.983
57	kotonoha (言の葉)	noun	4	1	36	54.327
58	domo (ども)	particle (conjective)	90	80	101	50.962
59	wakaru (別る)	verb	13	4	43	50.796
60	kage (影)	noun	27	10	53	49.697
61	momijiba (紅葉)	noun	1	1	30	49.542
62	tamoto (袂)	noun	2	0	29	47.738
63	ji (じ)	auxiliary verb	123	252	150	47.028
64	rashi (らし)	auxiliary verb	0	1	27	46.601
65	miru (見る)	verb	1299	644	457	46.534
66	no (野)	noun	15	7	43	46.268
67	uguisu (鶯)	noun	15	2	40	46.107
68	o (を)	particle (conjective)	45	85	84	45.911
69	se (瀬)	noun	2	1	29	45.639
70	yo (世)	noun	250	326	193	45.359
71	akikaze (秋風)	noun	0	2	27	44.136
72	koromo (衣)	noun	4	2	31	44.090
73	utsurou (移ろふ)	verb	9	4	36	43.776
74	ura (浦)	noun	13	2	37	43.342
75	ku (来)	verb	582	267	246	43.250
76	kumo (雲)	noun	27	5	46	43.206
77	beranari (べらなり)	auxiliary verb	0	1	25	42.945
78	ashihiki (足引き)	noun	0	0	23	42.796
78	shirakumo (白雲)	noun	0	0	23	42.796
80	furu (降る)	verb	162	46	103	42.708
81	ama (天)	noun	5	1	30	42.385
82	tsutsu (つつ)	particle (conjective)	315	57	143	42.233
83	shigure (時雨)	noun	9	3	34	41.635
84	na (名)	noun	71	34	72	41.500
85	kyoo (今日)	noun	83	112	98	41.339
86	waku (分く)	verb	45	23	59	40.984
87	yamabe (山辺)	noun	2	1	26	40.327
88	nobe (野辺)	noun	1	3	27	40.319
89	nu (ぬ)	auxiliary verb	1859	1131	620	40.195
90	aku (飽く)	verb	25	11	45	39.311
91	nashi (無し)	adjective	865	483	333	39.226
92	karakoromo (唐衣)	noun	0	0	21	39.075
92	shiratsuyu (白露)	noun	0	0	21	39.075
94	sora (空)	noun	70	21	65	38.799
95	moyu (燃ゆ)	verb	4	1	27	38.736
96	yume (夢)	noun	45	39	62	38.031
97	ou (生ふ)	verb	27	12	45	37.581
98	ma (間)	noun	58	25	61	37.286
99	tomo (とも)	particle (conjective)	49	156	96	37.207
100	goto (毎)	suffix	45	16	53	36.851

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
101	o (を)	particle (case)	3647	2178	1062	36.678
102	nuru (濡る)	verb	23	9	41	36.297
103	yoso (余所)	noun	8	18	38	35.981
104	harugasumi (春霞)	noun	0	1	21	35.650
105	kara (から)	particle (case)	40	10	47	34.523
106	yuki (雪)	noun	82	33	68	34.139
107	wabu (詫ぶ)	verb	51	21	54	33.557
108	shiranami (白波)	noun	0	1	19	32.013
109	kumoi (雲居)	noun	2	2	22	31.674
110	ka (香)	noun	17	3	32	31.620
111	kesa (今朝)	noun	10	7	30	30.952
111	ato (跡)	noun	15	2	30	30.952
113	nu (寝)	verb	118	38	76	30.901
114	wasuru (忘る)	verb	64	54	65	30.341
115	mo (も)	particle (bound)	4506	2584	1227	30.195
116	oki (沖)	noun	0	0	16	29.770
116	namidagawa (涙川)	noun	0	0	16	29.770
118	matsu (松)	noun	36	6	40	29.639
119	oosaka (あふさか)	noun	12	3	28	29.573
120	kanashi (悲し)	adjective	129	61	82	28.589
121	amanogawa (天の川)	noun	2	3	21	28.561
121	yoshino (吉野)	noun	4	1	21	28.561
123	chihayafuru (千早振る)	noun	1	0	17	28.385
124	ume (梅)	noun	27	5	35	28.318
125	niou (匂ふ)	verb	15	0	27	28.090
126	ayanashi (文無し)	adjective	0	0	15	27.910
126	shirayuki (白雪)	noun	0	0	15	27.910
126	hitsu (漬つ)	verb	0	0	15	27.910
129	koyu (越ゆ)	verb	22	11	35	27.792
130	dani (だに)	particle (adverbial)	156	197	118	27.587
131	ha (葉)	noun	35	9	39	27.503
132	fukashi (深し)	adjective	47	29	49	26.719
133	moga (もが)	particle (final)	17	9	31	26.376
134	shimo (霜)	noun	11	1	24	26.055
135	ne (嶺)	noun	0	0	14	26.049
136	ko (木)	noun	25	4	32	26.028
137	masaru (勝る)	verb	84	42	62	25.505
138	momiji (紅葉)	noun	24	9	33	25.248
139	tatsuta (竜田)	noun	5	0	19	25.227
140	nageki (嘆き)	noun	1	14	25	25.160
141	koe (声)	noun	167	37	81	25.100
142	ho (穂)	noun	3	1	18	24.866
143	ominaeshi (女郎花)	noun	9	1	22	24.757
143	furusato (古里)	noun	10	0	22	24.757
145	hanasusuki (花薄)	noun	2	1	17	24.613
145	chiyo (千代)	noun	3	0	17	24.613
147	tou (訪ふ)	verb	22	5	30	24.502
148	wakare (別れ)	noun	6	10	25	24.434
149	na (な)	particle (final)	32	135	71	24.250
150	toko (床)	noun	0	0	13	24.188

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
151	kazu (数)	noun	27	24	38	23.485
152	tsuki (月)	noun	177	76	90	23.452
153	oru (折る)	verb	52	22	45	23.217
154	keburi (煙)	noun	12	6	25	23.079
155	karu (刈る)	verb	11	3	23	22.999
156	miyako (都)	noun	9	16	28	22.976
157	ayamegusa (菖蒲草)	noun	1	0	14	22.966
157	hisakata (久方)	noun	1	0	14	22.966
157	hiru (干る)	verb	1	0	14	22.966
160	kusa (草)	noun	30	11	34	22.842
161	arashi (嵐)	noun	2	0	15	22.778
162	watatsuumi (わたつうみ)	noun	0	0	12	22.327
163	uramu (恨む)	verb	29	17	35	22.000
164	yamazato (山里)	noun	8	2	20	21.712
165	tare (誰)	pronoun	94	100	74	21.578
166	kawaku (乾く)	verb	2	2	16	21.511
167	kari (雁)	noun	8	1	19	21.057
168	shita (下)	noun	46	12	38	21.025
169	tsumu (摘む)	verb	6	5	20	20.894
170	koromode (衣手)	noun	0	0	11	20.467
170	tagitsu (滾つ)	verb	0	0	11	20.467
170	michinoku (陸奥)	noun	0	0	11	20.467
173	hikari (光)	noun	23	7	28	20.464
174	omou (思ふ)	verb	1892	876	517	20.318
175	shika (鹿)	noun	3	2	16	20.301
176	tsurashi (辛し)	adjective	22	20	32	20.135
177	shigeshi (繁し)	adjective	27	10	30	19.815
178	tsumoru (積もる)	verb	11	8	23	19.736
179	sae (さえ)	particle (adverbial)	114	61	67	19.647
180	oku (置く)	verb	154	121	89	19.586
181	chidori (千鳥)	noun	3	0	14	19.493
181	manima (随)	noun	3	0	14	19.493
183	mine (峰)	noun	9	6	21	19.463
184	itou (厭ふ)	verb	0	1	12	19.374
184	okuyama (奥山)	noun	1	0	12	19.374
186	tama (玉)	noun	13	29	31	19.010
187	karu (枯る)	verb	12	4	21	18.828
188	amagumo (天雲)	noun	0	0	10	18.606
188	sohotsu (そほつ)	verb	0	0	10	18.606
188	tama (魂)	noun	0	0	10	18.606
188	natsuku (懐く)	verb	0	0	10	18.606
188	watatsumi (海神)	noun	0	0	10	18.606
193	e (方)	noun	5	2	16	18.253
194	uu (植う)	verb	19	8	25	18.172
195	katami (形見)	noun	5	4	17	18.040
195	ada (徒)	noun	6	3	17	18.040
197	karu (離る)	verb	6	0	15	17.648
198	shiratama (白玉)	noun	1	0	11	17.585
198	yowa (夜半)	noun	1	0	11	17.585
200	suminoe (住の江)	noun	1	1	12	17.557

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
201	chigiru (契る)	verb	18	8	24	17.408
201	hama (浜)	noun	23	3	24	17.408
203	michi (道)	noun	69	51	51	17.406
204	seki (関)	noun	25	2	24	16.957
205	azusayumi (梓弓)	noun	0	0	9	16.745
205	itsuwari (偽り)	noun	0	0	9	16.745
205	utsusemi (空蝉)	noun	0	0	9	16.745
205	kara (殻)	noun	0	0	9	16.745
205	sasagani (細蟹)	noun	0	0	9	16.745
205	harusame (春雨)	noun	0	0	9	16.745
211	toshi (年)	noun	91	182	84	16.739
212	asashi (浅し)	adjective	5	2	15	16.716
213	ne (根)	noun	11	4	19	16.714
214	inochi (命)	noun	15	37	32	16.698
215	wasuregusa (忘れ草)	noun	2	2	13	16.566
216	somu (染む)	verb	11	3	18	15.962
217	kawaru (変はる)	verb	46	30	38	15.924
218	mu (む)	auxiliary verb	1234	2219	604	15.850
219	nodokeshi (長閑けし)	adjective	1	1	11	15.836
219	yorozuyo (万代)	noun	2	0	11	15.836
221	kami (神)	noun	34	39	37	15.751

List of Lowest Frequency Words for Short Poem in 15 literary works

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
1	tamau (給ふ: respect)	verb	2779	3915	3	-715.626
2	tari (たり)	auxiliary verb	3997	1205	62	-390.272
3	nado (など)	particle (adverbial)	2437	540	1	-318.302
4	oon (御)	prefix	995	1673	2	-279.958
5	ito (いと)	adverb	1910	1092	13	-278.335
6	te (て)	particle (conjective)	9482	4002	884	-229.902
7	su (す)	auxiliary verb	893	1320	21	-176.492
8	haberi (侍り)	verb	209	1284	2	-152.784
9	namu (なむ)	particle (bound)	539	885	2	-145.352
10	iu (言ふ)	verb	3244	717	168	-133.228
11	sasu (さす)	auxiliary verb	390	795	1	-123.607
12	moosu (申す)	verb	322	776	2	-110.317
13	imiji (いみじ)	adjective	605	448	1	-109.346
14	mairu (参る)	verb	474	554	1	-106.647
15	ko (是)	pronoun	848	1000	41	-105.491
16	domo (共)	suffix	699	279	2	-97.463
17	owashimasu (おはします)	verb	193	664	0	-93.083
18	tatematsuru (奉る)	verb	343	642	4	-91.825
19	nari (なり: declarative)	auxiliary verb	5409	4119	771	-90.367
20	notamau (宣ふ)	verb	504	260	0	-82.972
21	otoko (男)	noun	661	100	0	-82.646
22	yoo (様)	noun	463	279	0	-80.580
23	onna (女)	noun	600	134	0	-79.710
24	koto (事)	noun	1407	1792	174	-76.108
25	okashi (可笑し)	adjective	721	105	5	-72.470
26	tote (とて)	particle (case)	1253	365	54	-69.071
27	tokoro (所)	noun	542	288	7	-68.017
28	sari (然り)	verb	319	508	7	-67.713
29	oboyu (覚ゆ)	verb	459	237	2	-67.392
30	owasu (おはす)	verb	194	411	0	-65.690
31	obosu (思す)	verb	239	360	0	-65.038
32	kikoyu (聞こゆ)	verb	471	548	19	-63.639
33	dono (殿)	suffix	147	438	0	-63.517
34	ge (気)	suffix	604	178	10	-56.912
35	mi (御)	prefix	385	563	22	-52.628
36	saburau (侍ふ)	verb	158	322	0	-52.109
37	hodo (程)	noun	962	368	52	-48.714
38	so (其)	pronoun	418	407	18	-47.457
39	hitobito (人々)	noun	357	149	2	-47.299
40	nagon (納言)	noun	250	180	0	-46.678
41	otodo (大臣)	noun	128	293	0	-45.700
42	tono (殿)	noun	199	260	1	-45.494
43	raru (らる)	auxiliary verb	376	310	14	-40.912
44	gimi (君)	suffix	184	172	0	-38.641
45	yomu (読む)	verb	365	106	4	-38.513
46	miya (宮)	noun	221	200	2	-38.385
47	ni (に)	particle (conjective)	1292	331	89	-37.968
48	mikado (帝)	noun	108	233	0	-37.012
49	nichi (日)	noun	185	154	0	-36.795
50	tachi (達)	suffix	156	182	0	-36.687

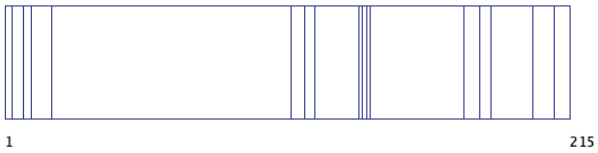
ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
51	rei (例)	noun	257	79	0	-36.469
52	sama (様)	noun	319	170	6	-36.067
53	monosu (ものす)	verb	184	142	0	-35.384
54	warau (笑ふ)	verb	275	49	0	-35.166
55	e (え)	adverb	256	222	6	-34.990
56	kuruma (車)	noun	252	132	2	-34.526
57	oosu (仰す)	verb	154	155	0	-33.538
58	mata (又)	conjunction	179	126	0	-33.103
59	kitanokata (北の方)	noun	200	102	0	-32.778
60	sate (扨)	conjunction	174	158	1	-31.983
61	in (院)	noun	58	234	0	-31.692
62	meri (めり)	auxiliary verb	289	407	22	-31.066
63	uta (歌)	noun	263	85	2	-30.788
64	juu (十)	noun	131	151	0	-30.606
65	kaku (書く)	verb	427	78	10	-30.583
66	iru (入る)	verb	633	161	30	-30.099
67	shooshoo (少将)	noun	213	63	0	-29.955
68	kore (此れ)	pronoun	270	389	21	-29.215
69	kashi (かし)	particle (final)	176	362	13	-29.118
70	iku (行く)	verb	242	59	1	-28.702
71	toru (取る)	verb	399	277	23	-28.317
72	sukoshi (少し)	adverb	192	100	1	-27.751
73	yoshi (良し)	adjective	216	223	8	-27.691
74	ichi (一)	noun	57	195	0	-27.349
75	sore (其れ)	pronoun	213	239	9	-27.293
76	medetashi (めでたし)	adjective	195	111	2	-26.450
77	idasu (出だす)	verb	183	47	0	-24.961
78	zu (ず)	verb	97	131	0	-24.744
79	ie (家)	noun	181	81	1	-24.588
80	tsuku (付く)	verb	452	200	25	-24.339
81	daijin (大臣)	noun	13	211	0	-24.309
82	chuujoo (中将)	noun	159	64	0	-24.201
83	miko (御子)	noun	86	131	0	-23.549
84	mae (前)	noun	259	132	8	-23.275
85	goran (御覧)	noun	86	128	0	-23.224
86	chuu (中)	prefix	123	91	0	-23.224
87	mono (者)	noun	148	99	1	-23.010
88	o (御)	prefix	168	102	2	-22.758
89	mina (皆)	noun	247	102	6	-22.608
90	kokochi (心地)	noun	349	124	14	-22.240
91	kami (守)	noun	163	76	1	-22.170
92	oboshimesu (思し召す)	verb	45	159	0	-22.138
93	musume (娘)	noun	90	112	0	-21.921
94	maji (まじ)	auxiliary verb	101	101	0	-21.921
95	kaerigoto (返り事)	noun	170	29	0	-21.595
96	moto (下)	noun	324	94	11	-21.397
97	mairasu (参らす)	verb	128	69	0	-21.378
98	su (す)	verb	2106	1404	319	-20.697
99	ayashi (怪し)	adjective	212	197	11	-20.625
100	tsukoomatsuru (仕る)	verb	77	113	0	-20.619

ranking	lexical items	parts of speech	narration	conversation	short poem	LLR values
101	dai (大)	prefix	99	90	0	-20.510
102	nijuu (二十)	noun	81	105	0	-20.184
103	ka (彼)	pronoun	161	143	5	-20.077
104	garu (がる)	suffix	154	29	0	-19.859
105	kaku (斯く)	adverb	321	449	40	-19.606
106	mata (又)	adverb	426	217	30	-19.079
107	tada (唯)	adverb	346	208	23	-18.969
108	fumi (文)	noun	314	103	13	-18.815
109	mesu (召す)	verb	97	75	0	-18.665
110	moozu (詣づ)	verb	104	68	0	-18.665
111	go (五)	noun	91	76	0	-18.122
112	nen (年)	noun	12	152	0	-17.796
113	notawamasu (宣はす)	verb	106	57	0	-17.688
114	gatsu (月)	noun	75	87	0	-17.579
115	ka (日)	suffix	250	58	7	-17.291
116	usu (失す)	verb	49	143	1	-17.258
117	tenjoo (殿上)	noun	98	59	0	-17.037
118	iru (居る)	verb	504	106	30	-16.836
119	ari (有り)	verb	2779	1620	435	-16.603
120	nonoshiru (罵る)	verb	134	19	0	-16.603
121	tamau (給ふ: humble)	verb	8	144	0	-16.494
122	san (三)	noun	56	96	0	-16.494
123	yoku (良く)	adverb	77	74	0	-16.385
124	ariku (歩く)	verb	140	43	1	-16.323
125	tadaima (只今)	noun	30	118	0	-16.060
126	itooshi (いとおし)	adjective	77	70	0	-15.951
127	monogatari (物語)	noun	117	30	0	-15.951
128	makaru (罷る)	verb	25	121	0	-15.843
129	kaeshi (返し)	noun	199	22	3	-15.737
130	sashi (差し)	prefix	240	63	8	-15.472
131	nikushi (憎し)	adjective	120	53	1	-15.286
132	emon (衛門)	noun	94	45	0	-15.083

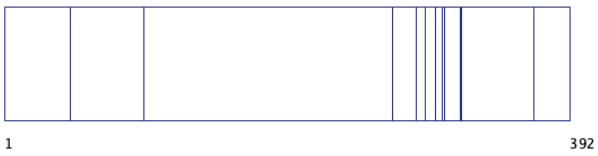
APPENDIX B

CONCORDANCE PLOTS FOR 54 CHAPTERS OF THE TALE OF GENJI

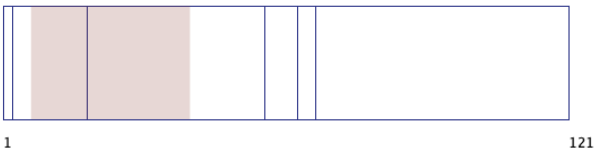
Chapter 1 Kiritsubo



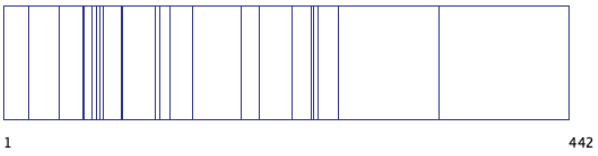
Chapter 2 Hahakigi



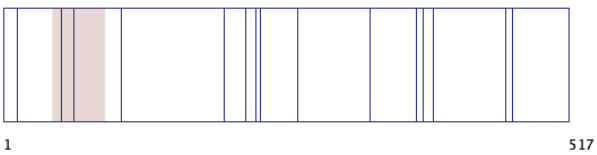
Chapter 3 Utsusemi



Chapter 4 Yugao

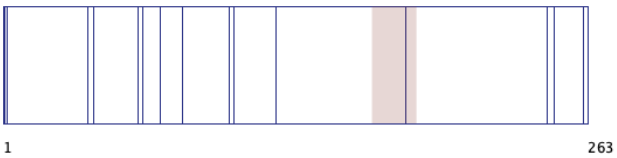


Chapter 5 Wakamursaki

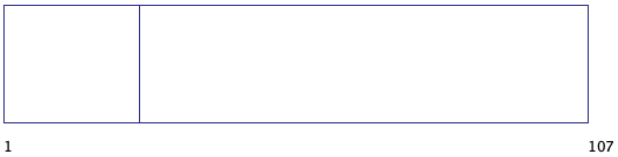


Chapter 6 Suetsumuhana

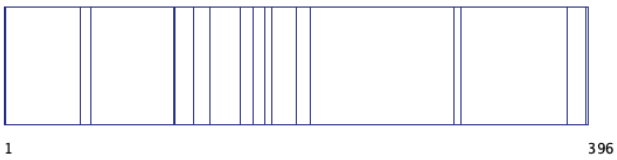
Chapter 7 Momiji no Ga



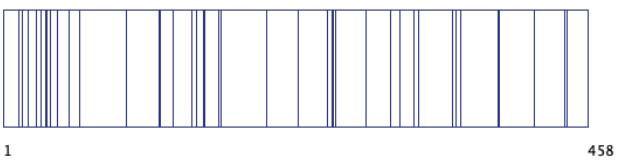
Chapter 8 Hana no En



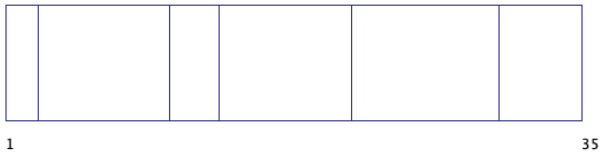
Chapter 9 Aoi



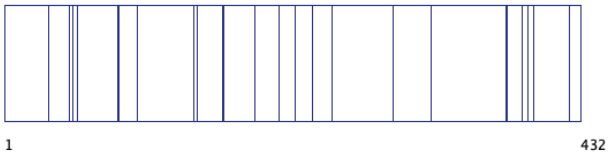
Chapter 10 Sakaki



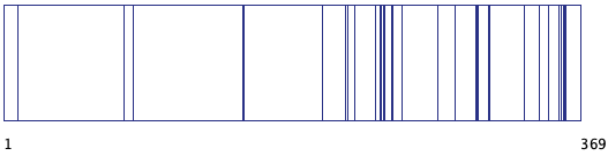
Chapter 11 Hana Chiru Sato



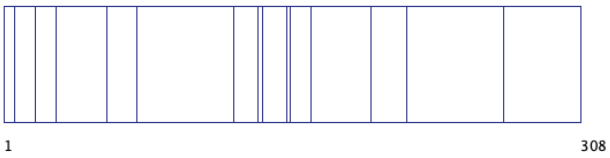
Chapter 12 Suma



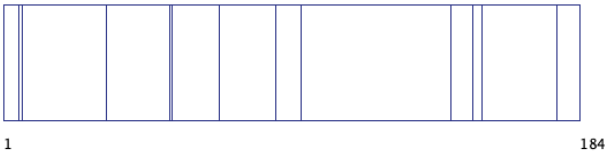
Chapter 13 Akashi



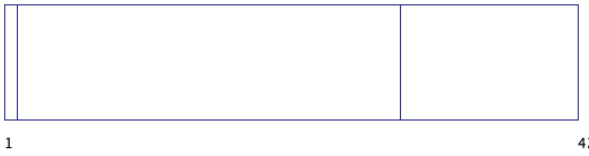
Chapter 14 Miotsukushi



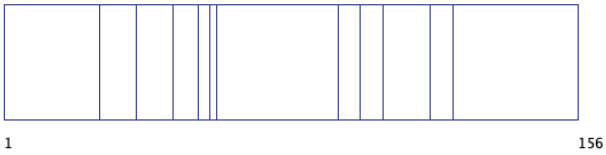
Chapter 15 Yomogyu



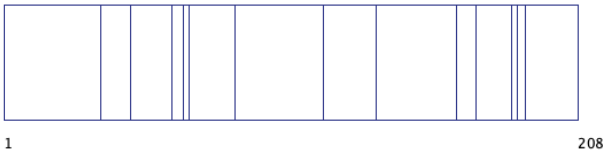
Chapter 16 Sekiya



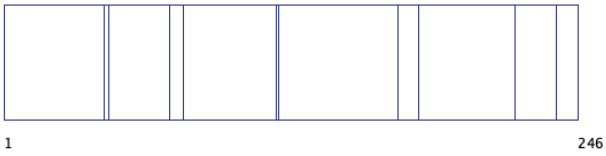
Chapter 17 E Awase



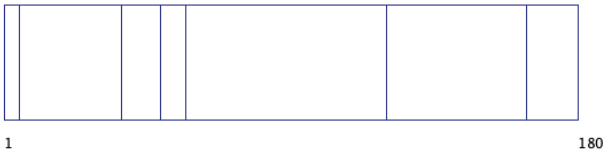
Chapter 18 Matsukaze



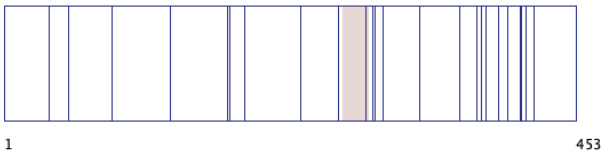
Chapter 19 Usugumo



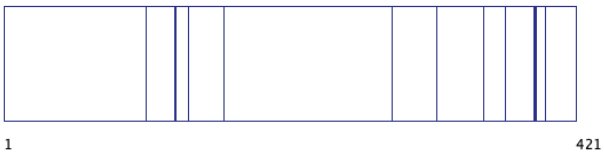
Chapter 20 Asagao



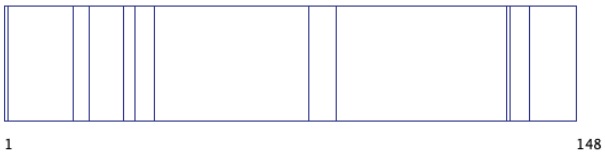
Chapter 21 Otome



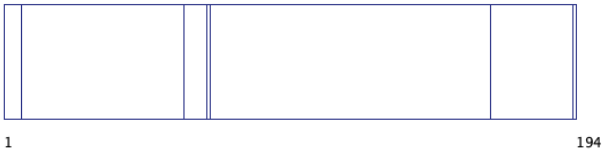
Chapter 22 Tamakazura



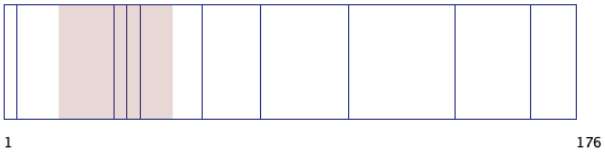
Chapter 23 Hatsune



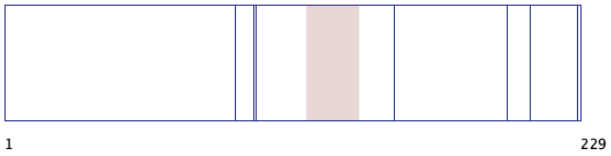
Chapter 24 Kocho



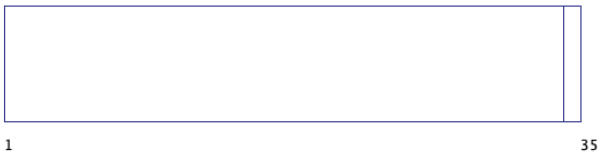
Chapter 25 Hotaru



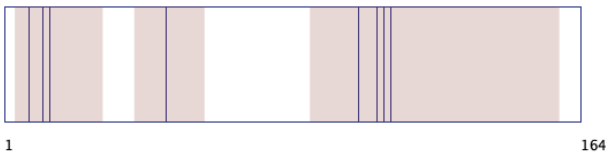
Chapter 26 Tokonatsu



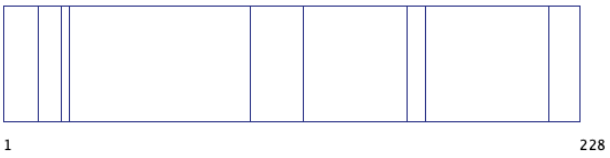
Chapter 27 Kagaribi



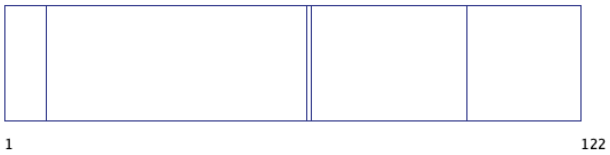
Chapter 28 Nowaki



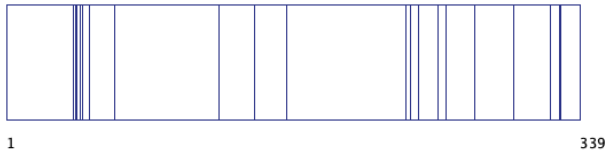
Chapter 29 Miyuki



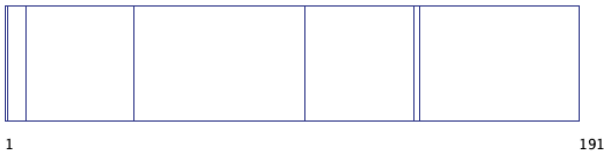
Chapter 30 Fujibakama



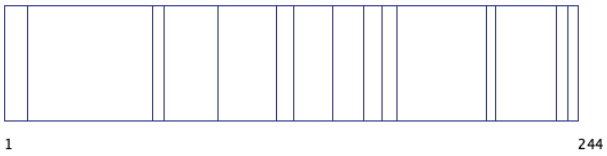
Chapter 31 Makibashira



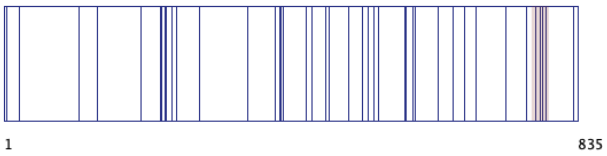
Chapter 32 Umegae



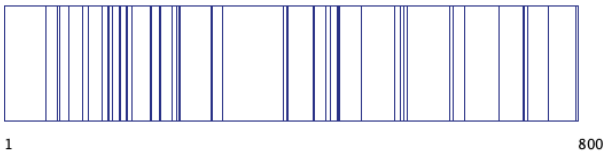
Chapter 33 Fuji no Uraba



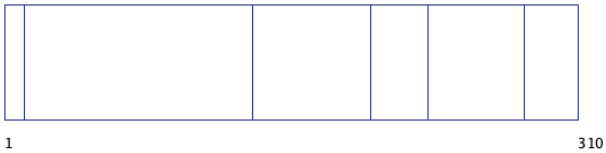
Chapter 34 Wakana: Jo



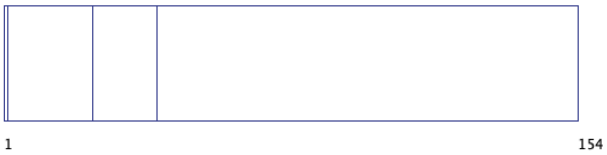
Chapter 35 Wakana: Ge



Chapter 36 Kashiwagi

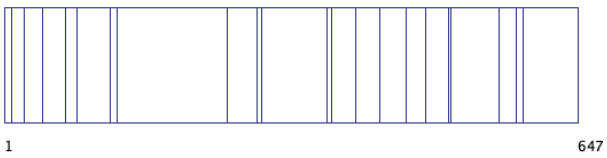


Chapter 37 Yokobue

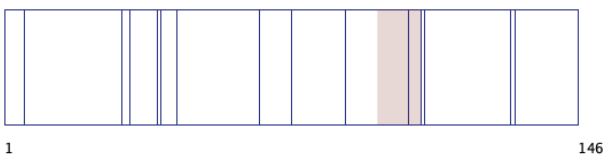


Chapter 38 Suzumushi

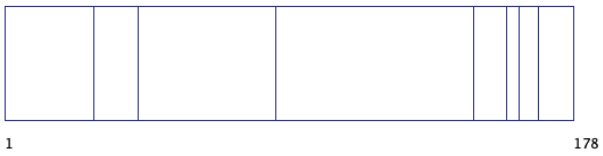
Chapter 39 Yugiri



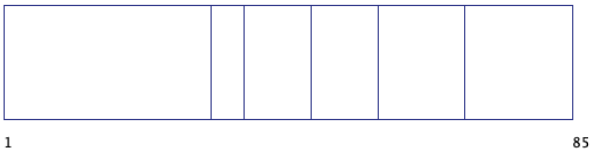
Chapter 40 Minori



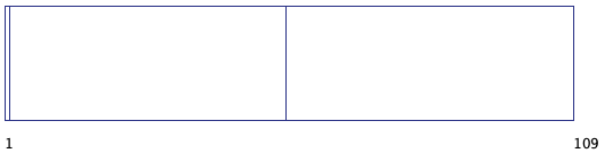
Chapter 41 Maboroshi



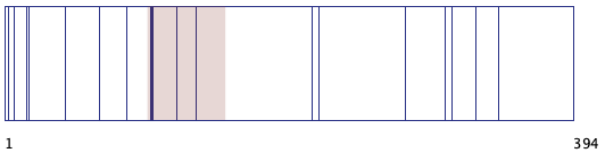
Chapter 42 Niou Hyobukyo



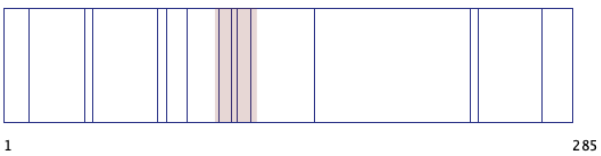
Chapter 43 Kobai



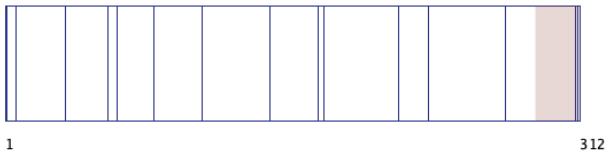
Chapter 44 Takekawa



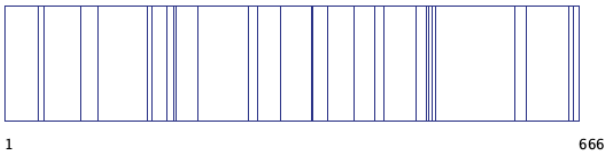
Chapter 45 Hashihime



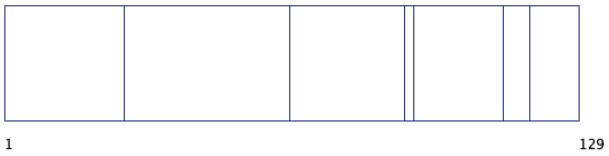
Chapter 46 Shi ga Moto



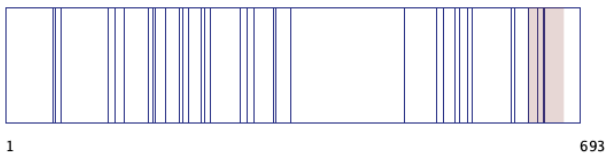
Chapter 47 Agemaki



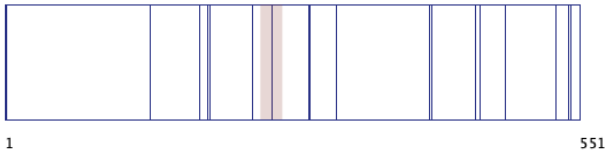
Chapter 48 Sawarabi



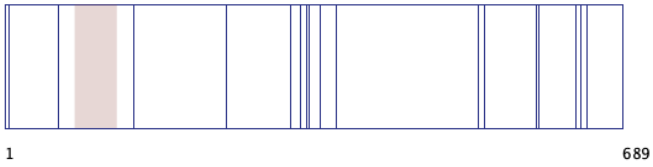
Chapter 49 Yadorigi



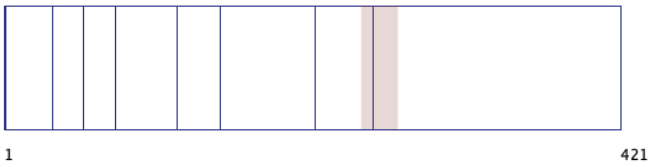
Chapter 50 Azumaya



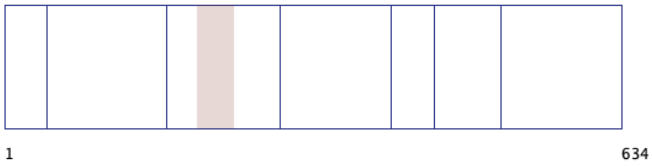
Chapter 51 Ukifune



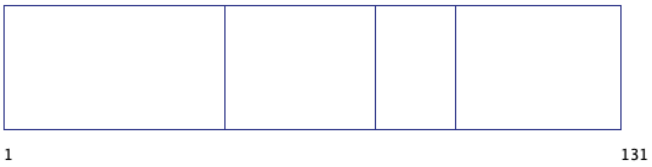
Chapter 52 Kagero



Chapter 53 Tenarai



Chapter 54 Yume no Ukihashi



REFERENCES

Akatsuka, N. (1978). Another look at *no*, *koto*, and *to*: Epistemology and complementizer choice in Japanese. In J. Hinds, & I. Howard (Eds.), *Problems in Japanese Syntax and Semantics* (pp. 178-212). Tokyo: Kaitakusha.

Akatsuka, N. (1979). Epistemology, Japanese Syntax, and Linguistic Theory. *Japanese Linguistics*, 6, 7-28.

Akiyama, K. (Ed.). (2000). *Oocho-go Jiten* [*Dictionary of Heian Lexical Items*]. Tokyo: The University of Tokyo Press.

Bal, M. (1985). *Narratology: Introduction to the Theory of Narrative*. Toronto: University of Toronto Press.

Banfield, A. (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction.* Boston: Routledge & Paul.

Bargen, D. (2015). *Mapping Courtship and Kinship in Classical Japanese: The Tale of Genji and its Predecessors*. Honolulu: University of Hawaii Press.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, *14(3)*, 275-311. doi:10.1075/ijcl.14.3.08bib

Biber, D. (2010). Corpus-Based and Corpus-driven Analyses of Language Variation and Use. In B. Heine, & H. Narrog (Eds.), *Oxford Handbook of Linguistic Analysis* (pp. 193-224). Oxford: Oxford University Press.

Biber, D. Gray, B. Staples, S. (2016). Contrasting the Grammatical Complexities of Conversation and Academic Writing: Implications for EAP Writing Development and Teaching. *Language in Focus Journal*, *2*, 1-18. doi: 10.1515/lifijsal-2016-0001

Fujii, S. (2001). *Heian Monogatari Jojutsu-ron* [*The Theory of Heian Narrative Literature*]. Tokyo: University of Tokyo Press.

Frellesvig, B. (2010). *A History of the Japanese Language*. Cambridge, New York: Cambridge University Press.

Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, 19 (1), 61-74.

Genette, G. (1980). *Narrative Discourse*. (J.E.Lewin Trans.). Ithaca; New York Cornell University Press.
(Original work published 1972)

Hamburger, K. (1973). *The Logic of Literature* (M. Rose, Trans.) Bloomington: Indiana University Press.
(Original work published 1957)

Higashihara, N. (1991). Kaisetsu: Kenkyu-shi to Tenbo [Commentary on the History of the Textual Analysis of *The Tale of Genji* and its Future Vision]. In K. Mitani & N. Higashihara (Eds.), *Genji monogatari: Katari to Hyogen* (pp. 242-257). Tokyo: Yuseido.

Hayashida, T. (1980). Kaimami no Bungei: Genji monogatari o Chushin ni shite [*Kaimami* in Literature: Focusing on *The Tale of Genji*]. In T. Hayashida (Ed.), *Genji monogatari no Hasso*. (pp.240-254). Tokyo: Ohfusha.
(Original work published 1976)

Hopper, P., & Thompson, S. (1980). Transitivity in grammar and discourse. *Language*, *56* (2), 251-299.

Ijima, M. (2011). *Chuukogo Kako Kanryoo Hyoogen no Kenkyuu* [*The Study on the Past Perfect Tense in Classical Japanese*]. Tokyo, Japan: Hituzi Syobo.

Imai, G. (2003). Oocho monogatari Kooseijo no Ichi-shuho: Kaimami ni tsuite [*Kaimami* as a Structural Technique in Heian Narrative Literature]. In Y. Imanishi, & N. Sakamoto (Eds.), *Imai Gene Chosaku-shu (Vol. 1)* (pp. 78-104). Tokyo: Kasamashoin.
(Original work published 1948)

Ikegami, Y. (1986). Nihongo no Katari no Tekusuto ni okeru Jisei no Tenkan ni tsuite [The Changing of Tense in Japanese Narratives]. *Kigoo Kenkyuu, 6*, 11-74.

Itoi, M. (1992). Monogatari Gengo no Hoo: Hyoogenshutai toshite no Katarite [Rules of Narrative Language: the Narrator as Expressive Subject]. In M. Itoi & T. Takahashi (Eds.), *Monogatari no Hoho* (pp. 21-34). Kyoto: Sekaishisosha.

Iwasaki, S. (1993). *Subjectivity in Grammar and Discourse: Theoretical considerations and a case study of Japanese spoken discourse*. Amsterdam: John Benjamins Publishing Company.

Iwasaki, S. (2013). *Japanese Revised Edition*. Amsterdam: John Benjamins Publishing Company.

Jinno, H. (2004). *Genji Monogatari to Wasei* [*Voice and expression in Genji Monogatari*]. Tokyo: Bensei Publishing.

Kai, M. (1980). *Bunshoo-hyoogen no Shikumi* [*The Mechanism of Textual Style*]. Tokyo: Ohfusha.

Katagiri, Y. (1999). *Uta-makura Uta-kotoba Jiten* [*Dictionary for Poetic Epithet Poetry Words*] (Revised Edition). Tokyo: Kasamashoin.

Kato, M. (2009). Honbun Kenkyuu to Ōshima-bon ni taisuru 15 no Gimon [15 Questions concerning the Textual Tradition of *The Tale of Genji* and the Ōshima-bon]. In The Association for the Study of Japanese Heian Literature, the West Branch (Eds.), *Ōshima-bon Genji monogatari no Saikento* (pp. 51-85). Osaka: Izumishoin.

Kawakami, N. (1976). Genji Monogatari ni okeru Kaimami no Kenkyuu [A Case Study on *Kaimami* Scenes in *The Tale of Genji*]. *Nihon Bungaku*, 46, 12-24.

Kilgariff, A. (2007). Comparison corpora. In W. Teubert, R. Krishnamurthy (Ed.), Corpus Linguistics: Critical Concepts in Linguistics (pp. 232-263). London; New York: Routledge.
(Original work published 2001)

Kinsui, S., Takayama, Y., Kinuhata, T., & Okazaki, T. (2011). *Shiriizu Nihongoshi3: Bunpooshi* [*The History of Grammar of Japanese Language*]. Tokyo: Iwanami Shoten.

Kitahara, Y., Suzuki, T., Takeca, T., Masubuchi, T. & Yamaguchi, Y. (1981). *Nihon Bunpoo Jiten* [*Encyclopedia of Japanese Grammar*]. Tokyo: Yuseido.

Kuno, S. (1973). *The Structure of the Japanese Language*. Cambridge: MIT Press.

Kuno, S. (1987). *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: University of Chicago Press.

Kurihara, Y. (1971). Ishiki no Enkinho [Consciousness Rendered in Perspective]. *Walpurgis 71*, 73-95. doi: 10.11501/4411375

Kuroda, S.-Y. (2014). Where epistemology, style, and grammar meet: A case study from Japanese. In S. Patron (Ed.), *Toward a Poetic Theory of Narration: Essays of S.-Y. Kuroda* (pp.38-59). Berlin, Germany: Mouton De Gruyter.
(Original work published 1973)

Kuroda, S.-Y. (2014). On grammar and narration. In S. Patron (Ed.), *Toward a Poetic Theory of Narration: Essays of S.-Y. Kuroda* (pp.60-70). Berlin, Germany: Mouton De Gruyter.
(Original work published 1974)

Kuroda, S.-Y. (2014). Reflections on the foundations of narrative theory, from a linguistic point of view. In S. Patron (Ed.), *Toward a Poetic Theory of Narration: Essays of S.-Y. Kuroda* (pp.71-101). Berlin, Germany: Mouton De Gruyter.
(Original work published 1975)

Leech, G. and Short, M. (2007). *Style in fiction. A linguistic introduction to English fictional prose*. Second edition. Harlow: Pearson.

Lyons, J. (1995). *Linguistic Semantics: An Introduction*. Cambridge: Cambridge University Press.

Matsushita, S. (2006). A Study of *The Tale of Genji* Focusing on Interior Monologue (Doctoral dissertation, The Australian National University, Canberra, Australia). Retrieved from http://digitalcollections.anu.edu.au/handle/1885/49400.

Mitamura, M. (1996). *Genji monogatari Kankaku no Ronri* [*The Tale of Genji*: Logic of Sensation]. Tokyo: Yuseido.

Mitani, K. (1989). *Monogatari Bungaku no Hoohoo* [*The Mechanism of Heian Narrative Literature*]. Tokyo: Yuseido

Mitani, K. (1994). *Genji monogatari no "Katari" to "Gensetsu"* [*"Story-telling" and "Discourse" in The Tale of Genji*]. Tokyo: Yuseido.

Mitani, K. (2002). *Genji Monogatari no gensetsu* [*Discourse of The Tale of Genji*]. Tokyo: Kanrin Shobo.

Murakami, F. (1998). Using Epistemic Modal Suffixes and Sensation/ Emotion Adjectives to Determine Narrative Distance in The Tale of Genji. *The Journal of the Association of Teachers of Japanese*, 32, 1-26. doi: 10.2307/489577

Murofushi, S. (1995). Kaimami no Hyogensei [Expressiveness of *Kaimami*]. In S. Murofushi (Ed.), *Ocho monogatari-shi no Kenkyu* (pp. 569-575). Tokyo: Kadokawa.
(Original work published 1983)

Nakano, K. (1969). Kodai monogatari no Dokusha no Mondai: Monogatari Ondokuron Hihan [Issues on Readers of the Heian Narrative Literature: A Critique of Monogatari Ondokuron]. In Nihon Bungaku Kenkyu Shiryo Kanko-kai (Eds.), *Genji monogatari I* (pp. 194-202). Tokyo: Yuseido.
(Original work published 1963)

Nakano, K. (1969). Genji monogatari no Soshiji to Monogatari Ondokuron [The Narrator's Comments and Monogatari Ondokuron in *The Tale of Genji*]. In Nihon Bungaku Kenkyu Shiryo Kanko-kai (Eds.), *Genji monogatari I* (pp. 203-212). Tokyo: Yuseido.
(Original work published 1964)

Narrog, H. (2012). *Modality, Subjectivity and Semantic Change: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Nishida, N. (1987). Keigo [Honorifics]. In A. Yamaguchi (Ed.), *Kokubunpoo Kooza* (pp.97-138). Tokyo: Meijishoin

Nishida, N. (2003). Keigo-shi to Gendai Keigo [The History of Honorifics and its Modern Usage]. In Y. Kikuchi (Ed.), Asakura Nihongo Koza (Revised Edition) (pp. 225-251). Tokyo. Yuseido.

Nishio, M. (1991). Genji monogatari ni okeru Taikenwaho (erlebte Rede) ni tsuite [Taikenwahoo (elebte Rede) in *The Tale of Genji*]. In N. Higashihara, & K. Mitani (Eds.), *Genji monogatari Katari to Hyogen* (pp. 28-39). Tokyo: Yuseido. (Original work published 1952)

Ogiso, T. (2015). Chuuko Wabun ni okeru Buntai betsu no Tokuchoogo [Lexical Items Showing Prominent Appearance in Different Styles in Classical Japanese]. In Y. Kondo, M. Tanaka, & T. Ogiso (Eds.), *Text Corpora and Japanese Historical Linguistics* (pp. 93-116). Tokyo, Japan: Hituzi Syobo.

Oda, M. (2015). *Jitsurei Shokai Koten Bunpoo Soran* [*The Complete Guide to the Grammar of Classical Japanese*]. Tokyo: Izumishoin.

Pascal, R. (1977). *The Dual Voice: Free Indirect Speech and Its Functioning in the Nineteenth-century European Novel*. Manchester: Manchester University Press.

Prince, G. (2003). A Dictionary of Narratology (Revised Edition). Lincoln; New York: University of Nebraska Press. (Original work published 1987)

Rimmon-Keenan, S. (1983). *Narrative Fiction: Contemporary Poetics*. London, New York: Methuen.

Saigo, N. (1983). *Genji monogatari o Yomu tameni* [*To read The Tale of Genji*]. Tokyo: Heibonsha.

Shibatani, M. (1990). *The Languages of Japan*. Cambridge: Cambridge University Press.

Shinohara, Y. (1973). Genji monogatari ni itaru Nozokimi no Keifu [The History of "Peering" up to *The Tale of Genji*]. *Kikan Bungaku Gengo*, 68, 56-67.

Short, M. (1996). *Exploring the Language of Poems, Plays and Prose*. London, New York: Longman.

Sinclair, J. (1991). *Corpus, Concordance, Collection*. Oxford: Oxford University Press.

Stinchecum, A. (1980). Who Tells the Tale?: 'Ukifune': A Study in Narrative Voice. *Monumenta Nipponica*, *35*, 375-403. doi: 10.2307/2384415

Sugiyama, Y. (1991). Genji monogatari no Katari no Shutai: Sono Kyokoo no Koozoo ni tsuite [Narrators in *The Tale of Genji*: The Structure of its Fiction]. In K. Mitani & N. Higashihara (Eds.), *Genji monogatari Katari to Hyoogen* (pp. 123-149). Tokyo. Yuseido.
(Original work published 1973)

Suzuki, K. (1969). Genji monogatari no Bunsho [The Text of *The Tale of Genji*]. *Kokubungaku: Kaishaku to Kansho*. 6. 34-36.

Takahashi, K. (1996). *Mechanisms of Tense Switch in Japanese Oral Narratives: (Re)enactment/ Recollection and Involvement/ Considerateness* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9620707)

Takahashi, T. (1991). *Monogatari to E no Enkinhoo* [*Point of view in Narrative Tales and Painting*]. Tokyo: Perikansha Publishing.

Tamagami, T. (1959). Genji monogatari no Keigo-hoo [The Usage of Honorifics in *The Tale of Genji*]. In A. Yamaguchi (Eds.), *Kooza Kaishaku to Bunpoo* (206-235). Tokyo: Meiji Shoin.

Tamagami, T. (1966). *Genji monogatari Hyoshaku: Bekkan 1* [*Commentary of The Tale of Genji 1*]. Tokyo: Kadokawa.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Tokieda, M. (1941). *Kokugo-gaku genron* [*Basic Theory of Japanese Language Theory*]. Tokyo: Iwanami.

Tokieda, M. (1950). *Nihon Bunpoo*. [*Japanese Grammar*]. Tokyo: Iwanami.

Traugott, E., & Dasher, R. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Watase, S. (2013). *Oochoo Jodooshi Kinooron: Anatanaru-ba, Wakukoozoo, Enkinhoo* [*The Functions of Auxiliary Verbs in Classical Japanese*]. Osaka: Izumishoin.

Yamada, Y. (1908). *Nihon Bunpooron* [*A Theory of Japanese Grammar*]. Tokyo: Hounkan.

Yamaoka, M. (2000). *"Katari" no Kigooron: Nichi-ei Hikaku Monogatari Bunseki* [*A Theory of Narratives: A Comparative Analysis of English and Japanese Narratives*]. Tokyo: Shohakusha.

Yoshikai, N. (2008). *"Kaimami"ru Genji Monogatari: Murasaki Shikibu no Shuhoo wo Kaiseki suru*[*"Peering" in The Tale of Genji: Analyzing Lady Murasaki's techniques*]: Tokyo: Kazamashoin.